STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 97963

TO: Michael Borin **Location: 12a01 / 12d01** **Wednesday, July 09, 2003** **Art Unit: 1631** **Phone: 305-4506** **Serial Number: 09 / 772538**	**From: Jan Delaval** **Location: Biotech-Chem Library** **CM1-1E07** **Phone: 308-4498** **jan.delaval@uspto.gov**

Search Notes

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

To: Jan Delaval

99163

RECEIVED

JUL - 2 2003

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

(STIC)

Requester's Full Name: M. BORIN Examiner #: 74104 Date: 07/02

Art Unit: 1631 Phone Number 30 5-4506 Serial Number: 09/772538

Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

For Sequence Searches Only *Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Please search claims 1 – 9.

For better understanding, I am attaching description of preferred embodiment

Thank you

M. Borin

****

STAFF USE ONLY

Searcher: _____

Searcher Phone #: 4478

Searcher Location: _____

Date Searcher Picked Up: 7/9/03

Date Completed: 7/9/03

Searcher Prep & Review Time: _____

Clerical Prep Time: 15

Online Time: ~90

Type of Search

NA Sequence (#) _____

AA Sequence (#) _____

Structure (#) _____

Bibliographic V

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN V

Dialog _____

Questel/Orbit _____

Dr.Link _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) _____

PTO-1590 (8-01)

Jean Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jean.delaval@uspto.gov

```
=> fil wpix
FILE 'WPIX' ENTERED AT 13:26:40 ON 09 JUL 2003
COPYRIGHT (C) 2003 THOMSON DERWENT

FILE LAST UPDATED:          7 JUL 2003      <20030707/UP>
MOST RECENT DERWENT UPDATE:     200343        <200343/DW>
DERWENT WORLD PATENTS INDEX SUBSCRIBER FILE, COVERS 1963 TO DATE

>>> NEW WEEKLY SDI FREQUENCY AVAILABLE --> see NEWS  <<<

>>> SLART (Simultaneous Left and Right Truncation) is now
    available in the /ABEX field. An additional search field
    /BIX is also provided which comprises both /BI and /ABEX <<<

>>> PATENT IMAGES AVAILABLE FOR PRINT AND DISPLAY <<<

>>> FOR DETAILS OF THE PATENTS COVERED IN CURRENT UPDATES,
    SEE http://www.derwent.com/dwpi/updates/dwpicov/index.html <<<

>>> FOR A COPY OF THE DERWENT WORLD PATENTS INDEX STN USER GUIDE,
    PLEASE VISIT:
 http://www.stn-international.de/training_center/patents/stn_guide.pdf <<<

>>> FOR INFORMATION ON ALL DERWENT WORLD PATENTS INDEX USER
    GUIDES, PLEASE VISIT:
    http://www.derwent.com/userguides/dwpi_guide.html <<<

=> d all abeq tech abex tot

L14  ANSWER 1 OF 7  WPIX   (C) 2003 THOMSON DERWENT
AN   2000-182459 [16]   WPIX
DNN  N2000-134601       DNC C2000-057101
TI   Identifying ligands for proteins, potentially useful as pharmaceuticals.
DC   B04 D16 J04 S03
IN   FROEMMEL, C; GOEDE, A; PREISSNER, R;
     FROMMEL, C
PA   (JERI-N) JERINI BIO TOOLS GMBH; (FROM-I) FROMMEL C; (GOED-I)
     GOEDE A; (PREI-I) PREISSNER R
CYC  21
PI   WO 2000004380 A1 20000127 (200016)* DE   20p    G01N033-53
        RW: AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE
         W: JP US
     DE 19831758   A1 20000203 (200018)             G01N033-68
     EP 1095272    A1 20010502 (200125)  DE         G01N033-53
        R: AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
     US 2002048776 A1 20020425 (200233)             G01N033-53      <--
ADT  WO 2000004380 A1 WO 1999-EP4951 19990713; DE 19831758 A1 DE 1998-19831758
     19980715; EP 1095272 A1 EP 1999-934689 19990713, WO 1999-EP4951 19990713;
     US 2002048776 A1 US 2001-772538 20010129
FDT  EP 1095272 A1 Based on WO 200004380
PRAI DE 1998-19831758 19980715
IC   ICM  G01N033-53; G01N033-68
     ICS  C07K001-00; C12Q001-00; G01N021-00; G01N033-542
AB   WO 200004380 A UPAB: 20000330
     NOVELTY - Determining ligands for proteins by determining the secondary
     structure elements of the ligand which form the ligand binding site, is
     new.
          DETAILED DESCRIPTION - Determining ligands (I) for proteins (II)
     comprises:
          (i) determining secondary structure elements (SSE) of (I) that form
     the ligand binding site;
          (ii) partitioning the molecular surface of (I) into surface patches
     (MSP);
```

(iii) identifying known areas (A) similar to the MSP that define the binding region, such that (A) possess a complementary neighbor element (B), co-ordinate transformation of (A) to (B) on a starting element at root mean square (rms) value below 2 Angstrom , and

(iv) estimating fitting of (I) according to local packing densities.

USE - The method is specifically used for determination of protein structures or for development of pharmaceuticals (claimed). Other applications for identified (I) are as diagnostic reagents, ligands for taste receptors (flavorings) and for affinity purification.

ADVANTAGE - The method determines (I) quickly and reliably.

Dwg.0/0

FS CPI EPI

FA AB

MC CPI: B11-C08; B12-K04E; D05-H09; D05-H10; D05-H13; J04-B01
EPI: S03-E14H4

TECH UPTX: 20000330

TECHNOLOGY FOCUS - BIOLOGY - Preferred Method: The outer surface of the secondary structure is determined, and outer surface elements that form contacts are selected as MSP. The similar known structures identified in the method are subjected to co-ordinate transformation, preferably to rms 1.5Angstrom, and then superimposed on the starting surface so that they lie on the atoms of the binding site. The best potential ligands are selected as lead compounds and/or are compared with known starting protein plus ligands.

Preferred ligands are determined as peptides of about 10 amino acids, then converted to peptidomimetics. (II) is an enzyme.

ABEX UPTX: 20000330

WIDER DISCLOSURE - Also disclosed is (I) determined by the new method and similar methods for identifying ligands that bind to RNA or DNA.

EXAMPLE - Starting from a binding site of an active subunit of the yeast proteosome, the secondary elements that define the binding site were determined. Five elements were detected, two of which formed the binding site. The exterior surface of these secondary elements was determined and the portion of the outer surface that made contacts (containing 12-22 atoms) was used to screen the DIP database for similar molecular surface patches (MSP). MSP with a minimum of 70% superimposition and root mean square value 1Angstrom were selected then superimposed on the starting surface. By co-ordinate transformation, reverse MSPs were formed, in the binding pocket, and assessed as potential ligands from their ability to fill the pocket and from whether the distances to atoms of the pocket were of correct size (local density). The best ligands were taken as lead compounds. A comparison of the 10 best ligands with a proteosome structure of an Archaebacterium, with ligand attached, showed that the main chain of one ligand was identical with the known inhibitor of this proteosome.

L14 ANSWER 2 OF 7 WPIX (C) 2003 THOMSON DERWENT

AN 1997-385296 [35] WPIX

CR 2001-606375 [55]; 2002-414103 [23]

DNC C1997-123570

TI Computer system containing crystallographic coordinates of human osteogenic protein-1 - used for design of analogues with agonist activity, potentially useful for treating osteoporosis etc..

DC B04

IN GRIFFITH, D; KECK, P; RUEGER, D C; SAMPATH, K T; CARLSON, W D; GRIFFITH, D L

PA (CREA-N) CREATIVE BIOMOLECULES INC; (UYBR-N) UNIV BRANDEIS

CYC 21

PI WO 9726277 A2 19970724 (199735)* EN 176p C07K014-51 <--
RW: AT BE CH DE DK ES FI FR GB GR IE IT LU MC NL PT SE
W: AU CA JP
AU 9722449 A 19970811 (199747) C07K014-51
EP 876401 A2 19981111 (199849) EN C07K014-51

R: AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
JP 2000501744 W 20000215 (200019) 188p C07K014-495
AU 725295 B 20001012 (200055) C07K014-51
AU 2000053497 A 20001102 (200062)# C07K014-51
ADT WO 9726277 A2 WO 1997-US1071 19970122; AU 9722449 A AU 1997-22449 19970122; EP 876401 A2 EP 1997-905604 19970122, WO 1997-US1071 19970122; JP 2000501744 W JP 1997-526303 19970122, WO 1997-US1071 19970122; AU 725295 B AU 1997-22449 19970122; AU 2000053497 A Div ex AU 1997-22449 19970122, AU 2000-53497 20000818
FDT AU 9722449 A Based on WO 9726277; EP 876401 A2 Based on WO 9726277; JP 2000501744 W Based on WO 9726277; AU 725295 B Previous Publ. AU 9722449, Based on WO 9726277; AU 2000053497 A Div ex AU 725295
PRAI US 1996-589552 19960122; AU 2000-53497 20000818
IC ICM C07K014-495; C07K014-51
ICS C07K001-00; C07K007-06; C07K007-08; G06F017-00; G06F017-30; G06F017-50
ICA C07B061-00
AB WO 9726277 A UPAB: 20020711
A novel computer system comprises: (a) a memory containing atomic X-ray crystallographic coordinates defining at least part of human osteogenic protein-1 (OP-1); and (b) a processor able to generate a molecular model having the three-dimensional (3D) shape representative of at least part of OP-1. Also new are: (1) production of a morphogenic analogue (I) with OP-1-like activity by: (i) identifying a candidate compound having a 3D shape corresponding to that of at least part of OP-1; then (ii) preparing this compound; (2) a method for preparing an analogue (II) that modulates an OP-1 mediated biological effect by using a similar process to (1) to identify a compound with the 3D shape and solvent-accessible surface (SAS) corresponding to those of OP-1; and (3) compounds identified by the method.

USE - The computer system is used in the method to define potential therapeutic agents. Typical applications of these are in proliferation of progenitor cells and regeneration of damaged or diseased tissue (e.g. osteoporosis or other bone-remodelling diseases) but more generally in any situation where mimicking or agonism of OP-1 is required.

ADVANTAGE - The analogues identified have, compared with native OP-1, better stability and/or solubility under physiological conditions; improved tissue target specificity; better biodistribution and/or reduced clearance rates. They may also lack an epitope or region recognised by a cellular scavenging protein, or altered receptor binding characteristics.
Dwg.0/16
FS CPI
FA AB
MC CPI: B04-N02; B11-C09; B14-N01

L14 ANSWER 3 OF 7 WPIX (C) 2003 THOMSON DERWENT
AN 1996-433261 [43] WPIX
CR 1994-358467 [44]; 1998-052548 [05]
DNN N1996-365112
TI Model protein with variable regions three-dimensional structure computer modelling - receiving relative positional information contg hi and phi angle values between pairs of amino acids followed by establishing position for first amino acid in variable region.
DC S03 T01
IN SRINIVASAN, S; SUDARSANAM, P
PA (IMMV) IMMUNEX CORP
CYC 1
PI US 5557535 A 19960917 (199643)* 21p G01N037-00 <--
ADT US 5557535 A CIP of US 1993-55050 19930428, US 1994-234812 19940428
FDT US 5557535 A CIP of US 5453397
PRAI US 1994-234812 19940428; US 1993-55050 19930428
IC ICM G01N037-00
AB US 5557535 A UPAB: 19990511

The method involves receiving relative positional information between pairs of amino acids, followed by establishing a position for a first amino acid of the variable region. For each amino acid pair in the variable region, it requires generating a model position for the amino acids based on the received relative positional information for the pair of amino acids. The information includes hi and phi angle values between pairs of amino acids.

For each combination of pairs of amino acids in variable regions of a family of proteins, the method entails collecting the hi and phi angle values for each pair of amino acids and while the step of generating a model position bases the model position on one of the collected hi and phi angle values.

USE/ADVANTAGE - For modelling 3D structure of protein and sequence alignment between protein to be modelled and template protein. Minimizes short contacts and allows modelling proteins with weak sequence identity with template protein.

Dwg.2/10

FS EPI
FA AB; GI
MC EPI: S03-E14H5; T01-J15X

L14 ANSWER 4 OF 7 WPIX (C) 2003 THOMSON DERWENT
AN 1996-139250 [14] WPIX
DNN N1996-116692
TI Vaccine and drug design - using coordinate axes with origin in binding site of receptor molecule and grid with mesh of resolution sufficient for low-energy amino acid locations and finding min. energy interaction energy position between amino acid and site.
DC S05 T01
IN CORNETTE, J L; DELISI, C; ROSENFELD, R; SEZERMAN, U; VAJDA, S
PA (UYBO-N) UNIV BOSTON
CYC 1
PI US 5495423 A 19960227 (199614)* 22p G06F017-50 <--
ADT US 5495423 A US 1993-142597 19931025
PRAI US 1993-142597 19931025
IC ICM G06F017-50
AB US 5495423 A UPAB: 19960405
A method for computing the conformation and location that a protein fragment will obtain in binding to the active site of a receptor is provided. The invention relates to both the computation of conformation and location of natural ligands within an active site and the design of artificial ligands with useful binding characteristics.
FS EPI
FA AB
MC EPI: S05-X; T01-J15X

L14 ANSWER 5 OF 7 WPIX (C) 2003 THOMSON DERWENT
AN 1993-351641 [44] WPIX
CR 1989-294634 [41]
DNN N1993-271247 DNC C1993-156070
TI New biologically active peptide with stabilised three-dimensional structure - mimics active site protein, useful e.g. as HIV or malaria vaccines, resistant to temp. and solvent changes.
DC B04 D16 S03
IN ARRHENIUS, T; CABEZA, E; CHIANG, L; SATTERTHWAIT, A C; CABEZAS, E
PA (SCRI) SCRIPPS RES INST
CYC 22
PI WO 9321206 A1 19931028 (199344)* EN 128p C07K001-00 <--
RW: AT BE CH DE DK ES FR GB GR IE IT LU MC NL PT SE
W: AU CA FI JP NO
AU 9339718 A 19931118 (199410) C07K001-00
US 5807979 A 19980915 (199844) C07K007-54
ADT WO 9321206 A1 WO 1993-US3032 19930331; AU 9339718 A AU 1993-39718

19930331; US 5807979 A Cont of US 1988-179160 19880408, Cont of US 1990-607645 19901029, CIP of US 1991-746064 19910812, Cont of US 1992-866040 19920408, Cont of US 1994-224059 19940407, US 1995-456424 19950601

FDT AU 9339718 A Based on WO 9321206

PRAI US 1993-33883 19930319; US 1992-866040 19920408; US 1988-179160 19880408; US 1990-607645 19901029; US 1991-746064 19910812; US 1994-224059 19940407; US 1995-456424 19950601

REP 5.Jnl.Ref; US 4642334; US 4777127

IC ICM C07K001-00; C07K007-54
ICS A61K037-02; A61K039-00; A61K039-395; C07K014-155; C07K015-00; C12N015-00; G01N033-53

AB WO 9321206 A UPAB: 19981104
New biologically active peptide (I) has a three-dimensionally stabilised configuration which mimics that of an active site of a natural biologically active protein. The configuration is stabilised by a covalently bound linking gp. which connects non-adjoining aminoacid residues in the peptide.

Partic. the linking gp. (which replaces an H bond) forms one of the structures (I), (II) or (III) of R = H or 1-6C alkyl; R1, R2 and R3 = H or 1-6C alkyl; R3 may also be a chain of 1-2000 aminoacids; R4 = any atom or gp. with the required configuration; m = 0-5; aa = aminoacid residue; n = 1-2000, esp. 3-30.

Pref. the aminoacid sequence and 3-D configuration of the active site of natural protein is determined, then a hydrogen bond identified which stabilises the structure. A peptide is synthesis (by standard methods) having (practically) the same sequence as the active site, then a covalent link is formed to join the atoms involves in the stabilising H bond. Partic. the covalent bond is formed by reaction with a divalent linker. More than one hydrogen bond can be replaced.

USE/ADVANTAGE - (I) are stabilised against changes in temp. and solvent, and against natural degradation processes, but retain the activity of (much larger) peptides. They can mimic a wide variety of pharmaceuticals, hormones, diagnostic reagents, vaccines, catalysts etc., e.g. epidermal growth factor, and HIV or malaria peptides. Configurationally restricted peptides should provide a more efficient immune response, producing antibodies with a binding pocket which closely complements the native protein surface.

Dwg.0/6

FS CPI EPI

FA AB; GI; DCN

MC CPI: B02-V02; B04-C01; B12-B03; B12-K04; D05-H07
EPI: S03-E14H4

L14 ANSWER 6 OF 7 WPIX (C) 2003 THOMSON DERWENT

AN 1993-045645 [05] WPIX

DNN N1993-034929 DNC C1993-020663

TI Characterising the three-dimensional structure of a protein - by analysing aminoacid residue positions and comparing with known protein structures.

DC B04 D16 T01

IN BOWIE, J U; EISENBERG, D; LUTHY, R

PA (REGC) UNIV CALIFORNIA

CYC 18

PI WO 9301484 A1 19930121 (199305)* EN 56p G06F015-20 <--
RW: AT BE CH DE DK ES FR GB GR IT LU MC NL SE
W: AU CA JP
AU 9224082 A 19930211 (199321) G06F015-20
US 5436850 A 19950725 (199535) 23p G06F019-00

ADT WO 9301484 A1 WO 1992-US5773 19920710; AU 9224082 A AU 1992-24082 19920710; US 5436850 A Cont of US 1991-728640 19910711, US 1994-218685 19940328

FDT AU 9224082 A Based on WO 9301484

PRAI US 1991-728640 19910711; US 1994-218685 19940328

REP US 4704692; US 4717653; US 4853871; US 4881175; US 4908773; US 4939666; US 4946778; US 4976958; US 5087558
IC ICM G06F015-20; G06F019-00
ICS C12N015-00; C12Q001-68
AB WO 9301484 A UPAB: 19931119
Characterising the 3-dimensional(3-D) structure of a protein, comprises (a) determining, from the 3-D structure of the protein, values for a structural properties P1,P2....Pn for each amino acid residue position of the protein, (b) assigning each residue of the protein to one environment class based upon the values for the n structural properties P1,P2 ....Pn for the residue, thereby generating a 1-dimensional environment string comprising the environment class of each residue in the 3-D protein structure.
USE/ADVANTAGE - Permit the assignment of many amino acid sequences to known 3-D structures. Used partic. for screening structural analogues of a known protein sequence. The 3-D compatibility searches are able to detect structural relationships that may not be apparent by sequence similarity.
1/5
Dwg.1/5
FS CPI EPI
FA AB; GI
MC CPI: B04-B04A; B12-K04; D05-H09
EPI: T01-J10B2
ABEQ US 5436850 A UPAB: 19950905
The three-dimensional structure of a protein is characterised by determining values for n structural properties P1-Pn for each amino acid residue, and assigning each residue to one of a number of environmental classes based on the values to generate a one-dimensional environment string comprising the class of each residue. The data generated are input into a programmed computer which compares them to a database of other proteins of known structure and outputs analogous structures. The properties pref. include the total area of a residue side-chain buried by other protein atoms inaccessible to solvent, the fraction of the side-chain area covered by polar atoms or water, and the local secondary structure.
USE/ADVANTAGE - Partic. for identifying protein sequences which fold into a known three-dimensional structure. Relates a one-dimensional target sequence directly to known three-dimensional structures and effectively utilises information about the accommodation of sequence changes inherent in known structures.
Dwg.1/8

L14 ANSWER 7 OF 7 WPIX (C) 2003 THOMSON DERWENT
AN 1991-340015 [46] WPIX
DNN N1991-260465 DNC C1991-146838
TI 3-dimensional protein structure determining method - processing full sequence of aminoacid residues of protein using microprocessor for simulating folding of protein.
DC J04 S03 T01
IN SKOLNICK, J; KOLINSKI, A
PA (SCRI) SCRIPPS CLINIC & RES FOUND; (SCRI-N) SCRIPPS CLINIC & RE
CYC 20
PI WO 9116683 A 19911031 (199146)* <--
RW: AT BE CH DE DK ES FR GB GR IT LU NL SE
W: AU CA FI JP NO
AU 9178837 A 19911111 (199207)
JP 05501324 W 19930311 (199315) 43p G06F015-20
PT 97480 A 19930531 (199325) G06F007-00
US 5265030 A 19931123 (199348) 74p G06F015-60
ADT JP 05501324 W JP 1991-509821 19910423, WO 1991-US2786 19910423; PT 97480 A PT 1991-97480 19910424; US 5265030 A Cont of US 1990-513918 19900424, Cont of US 1991-803678 19911203, US 1992-932282 19920819
FDT JP 05501324 W Based on WO 9116683

PRAI US 1990-513918 19900424
REP US 4704692; US 4853871; US 4881175; US 4908773; US 4939666; US 4985827
IC ICM G06F007-00; G06F015-20; G06F015-60
ICS G01N033-68; G06F015-42; G06F015-46
AB WO 9116683 A UPAB: 19931220
The method determines with a machine a three-dimensional structure of a protein or portion thereof including sidechains is claimed. A sequence of amino acid residues whose native tertiary structure is to be determined and local conformation preferences for respective residues of the sequence are specified.

A temp. is also specified, and a repreesentation of an unfolded chain of the residues in three dimensions is automatically generated. Folding of the chain and interations between all pairs of sidechains are simulated, and the representation of the tertiary structure displayed.

ADVANTAGE - Simulation of protein folding and prediction of tertiary structure are not only performed with greater success and accomplished faster than by many existing methods, but simulation itself becomes more manageable (tractable). @(109pp Dwg.No.3/16)@
FS CPI EPI
FA AB; GI
MC CPI: J04-C
EPI: S03-E14H1; T01-J07
ABEQ JP 05501324 W UPAB: 19930928
Determining three-dimensional structure of protein or portion including side chains, using machine is claimed. Sequence of amino acid residues whose native tertiary structure is to be determined and local conformation preferences for respective residues of the sequence are specified.

Temp. is also specified, and representation of unfolded chain of the residues in three dimensions is automatically generated. Folding of the chain and interactions between all pairs of side chains are simulated, and the representation of the tertiary structure displayed.

ADVANTAGE - Simulation of protein folding and prediction of tertiary structure are not only performed with greater success and accomplished faster than by many existing methods, but simulation itself becomes more manageable (tractable)
ABEQ US 5265030 A UPAB: 19940120
Three-dimensional structures of globular proteins are determined by a computer-based system which is based on a Monte Carlo dynamics technique with asymmetric Metropolis sampling criterion. A sequence of amino acid residues of a protein is specified and a 210 lattice structure is created for each amino acid of the protein. An unfolded conformation consisting of alpha-C backbone and side chains is represented spatially and successive likely tertiary conformations are selected, from which the lowest total-free-energy conformation is chosen. Coordinate set is created for display.

ADVANTAGE - Fast method of determining protein structures.
Dwg.15B/16

=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 13:26:57 ON 09 JUL 2003

FILE COVERS 1907 - 9 Jul 2003 VOL 139 ISS 2
FILE LAST UPDATED: 8 Jul 2003 (20030708/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 14:16:12 ON 09 JUL 2003

FILE COVERS 1907 - 9 Jul 2003 VOL 139 ISS 2
FILE LAST UPDATED: 8 Jul 2003 (20030708/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=> d all tot

L86 ANSWER 1 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **2001:538255** HCAPLUS
DN **135:132768**
TI **Computational** methods for generation of synthetic **ligands** based on the three dimensional **structure** of thyroid hormone receptor
IN Scanlan, Thomas S.; Baxter, John D.; Fletterick, Robert J.; Wagner, Richard L.; Kushner, Peter J.; Apriletti, James J.; West, Brian L.; Shiau, Andrew K.
PA Regents of the University of California, USA
SO U.S., 268 pp., Cont.-in-part of U.S. 6,236,946.
CODEN: USXXAM
DT Patent
LA English
IC ICM G06F019-00
ICS G06F017-00; C07G014-00
NCL 702022000
CC 2-7 (Mammalian Hormones)
Section cross-reference(s): 3, 6
FAN.CNT 3

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 6266622	B1	20010724	US 1997-980115	19971126 <--
	CA 2240024	AA	19970619	CA 1996-2240024	19961213 <--
	US 6236946	B1	20010522	US 1996-764870	19961213 <--
	CA 2314096	AA	19990603	CA 1998-2314096	19981125 <--
	WO 9926966	A2	19990603	WO 1998-US25296	19981125 <--

WO 9926966 A3 20000120
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
AU 9917999 A1 19990615 AU 1999-17999 19981125 <--
EP 1034184 A2 20000913 EP 1998-962849 19981125 <--
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI
BR 9815061 A 20011120 BR 1998-15061 19981125 <--
JP 2001524489 T2 20011204 JP 2000-522122 19981125 <--
PRAI US 1995-8540P P 19951213 <--
US 1995-8543P P 19951213 <--
US 1995-8606P P 19951214 <--
US 1996-764870 A2 19961213 <--
US 1997-980115 A 19971126 <--
WO 1998-US25296 W 19981125 <--
OS MARPAT 135:132768

AB The present invention provides new methods, particularly **computational** methods, and compns. for the generation of nuclear receptor synthetic **ligands** based on the three dimensional **structure** of nuclear receptors, particularly the thyroid receptor (TR). Also provided are crystals, nuclear receptor synthetic **ligands**, and related methods. The present invention provides for crystals of TR **ligand binding** domains with a **ligand** bound to the **ligand binding** domain (LBD), which provide excellent at. resoln. of the **amino acids** that interact with TR **ligand**, esp. thyroid receptor **ligands**. The three dimensional **model** of a TR LBD with a **ligand** bound reveals a previously unknown **structure** for nuclear receptors and shows that the **ligand** is bound in a water inaccessible **binding** cavity of the **ligand binding** domain of the TR. The present invention also includes a method for identifying a compd. capable of selectively modulating the activity of a TR isoform. Further included is a method for identifying agonist or antagonist **ligands** of a TR using the at. **coordinates** of a LBD in conjunction with a **computerized modeling** system. Also provided is a method of identifying a compd. that selectively modulates the activity of one type of nuclear receptor compared to other nuclear hormone receptors. Another aspect of the invention is a method for increasing the receptor selectivity of a compd. for a particular type of nuclear receptor. The invention finds use in the selection and characterization of **peptide**, peptidomimetic or synthetic compds. identified by the methods of the invention, particularly new lead compds. useful in treating disorders related to nuclear receptor-based deficiencies, including TR-related disorders.

ST thyroid hormone receptor **structure** synthetic **ligand**

IT **Protein motifs**
(LBD (**ligand-binding** domain); **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT **Computer application**
Drug **screening**
Protein sequences
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT Thyroid hormone receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); BUU

(Biological use, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process); USES (Uses)
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT Hormone receptors
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT **Conformation**
Crystallization
Secondary structure
(**protein**; **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT **Structure**-activity relationship
(thyroid hormone; **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT Thyroid hormone receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process); USES (Uses)
(.alpha.; **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT Thyroid hormone receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process); USES (Uses)
(.beta.; **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT 121939-99-9 351904-91-1
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(**amino acid** sequence; **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT 51-24-1, triac
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)
(complexed with TR-.beta.; **computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT 51-23-0P 140396-69-6P 192766-75-9P 192766-76-0P 192766-77-1P 192766-78-2P 192766-79-3P 192766-80-6P 350855-16-2P
RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses)
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT 178877-78-6P
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); RCT (Reactant); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); RACT (Reactant or reagent); USES (Uses)
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)

IT 1596-67-4, thyronine
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)
IT 6893-02-3 91969-75-4 192766-82-8 350855-17-3
RL: RCT (Reactant); RACT (Reactant or reagent)
(**computational** methods for generation of synthetic **ligands** based on three dimensional **structure** of thyroid hormone receptor)
IT 351909-92-7
RL: PRP (Properties)
(unclaimed nucleotide sequence; **computational** methods for generation of synthetic **ligands** based on the three dimensional **structure** of thyroid hormone receptor)
IT 100785-35-1 111518-64-0 112540-09-7 127187-77-3 130038-23-2, RNA formation factor RXR (human clone .lambda.XR3-1 isoform .alpha. reduced) 149024-37-3 150474-29-6 170139-50-1 351909-86-9 351909-87-0 351909-88-1 351909-89-2 351909-90-5 351909-91-6
RL: PRP (Properties)
(unclaimed **protein** sequence; **computational** methods for generation of synthetic **ligands** based on the three dimensional **structure** of thyroid hormone receptor)
RE.CNT 119 THERE ARE 119 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Agrafiotis; US 5463564 1995
(2) Andrea, T; J Med Chem 1979, V22, P221 HCAPLUS
(3) Andrews; US 4741897 1988 HCAPLUS
(4) Anon; EP 335628 1989 HCAPLUS
(5) Anon; WO 9721993 1997 HCAPLUS
(6) Anon; WO 9807435 1998 HCAPLUS
(7) Anon; WO 9857919 1998 HCAPLUS
(8) Apriletti, J; J Biol Chem 1988, V263, P9409 HCAPLUS
(9) Apriletti, J; Protein Expression and Purification 1995, V6, P363 HCAPLUS
(10) Au-Fliegner; Mol Cell Biol 1993, V13, P5725 HCAPLUS
(11) Baniahmad, A; Mol Cell Biol 1995, V15, P76 HCAPLUS
(12) Barettino, D; Embo J 1994, V13, P3039 HCAPLUS
(13) Barker; Ann N Y Acad Sci 1960, V86, P545 HCAPLUS
(14) Beck-Peccoz, P; J Clin Endocrinol Metab 1994, V78, P990 MEDLINE
(15) Bhat, M; Biochem Biophys Res Commun 1995, V210, P464 HCAPLUS
(16) Blake, C; J Mol Biol 1978, V121, P339 HCAPLUS
(17) Blake, C; Nature 1977, V268, P115 HCAPLUS
(18) Bolger; Journal of Biological Chemistry 1980, V255(21), P10271 HCAPLUS
(19) Bourguet, W; Nature 1995, V375, P377 HCAPLUS
(20) Brent, G; N Engl J Med 1994, V331, P847 HCAPLUS
(21) Brunger, A; Science 1987, V235, P458
(22) Casanova, J; Mol Cell Biol 1994, V14, P5756 HCAPLUS
(23) Cavailles, V; Embo, J 1995, V14, P3741 HCAPLUS
(24) Chang, K; Proc Natl Acad Sci USA 1997, V94(17), P9040 HCAPLUS
(25) Chiellini; Chemistry and Biology 1998, V5(6), P299 HCAPLUS
(26) Chin; US 5284999 1994 HCAPLUS
(27) Collaborative Computational Project; Acta Crystallogr 1994, VD50(4), P760
(28) Collingwood, T; Mol Endocrinol 1994, V8, P1262 HCAPLUS
(29) Cowan, S; Acta Crystallogr A 1991, V47, P110
(30) Cowtan, K; Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 1994, V31, P34
(31) Crowe; Methods in Molecular Biology 1994, V31, P371 HCAPLUS
(32) Damm, K; Proc Natl Acad Sci USA 1993, V90, P10668 HCAPLUS
(33) Danielian, P; Embo J 1992, V11, P1025 HCAPLUS
(34) Davies; US 5322933 1994 HCAPLUS
(35) Dawson; US 5466861 1995 HCAPLUS
(36) Degroot; US 5438126 1995 HCAPLUS
(37) Dietrich, S; J Med Chem 1977, V20, P863 HCAPLUS

(38) Durand, B; Embo J 1994, V13, P5370 HCAPLUS
(39) Ellis; US 4766121 1988 HCAPLUS
(40) Ellis; US 4826876 1989 HCAPLUS
(41) Ellis; US 4910305 1990 HCAPLUS
(42) Emmett; US 5061798 1991 HCAPLUS
(43) Evans; US 5171671 1992 HCAPLUS
(44) Evans; US 5312732 1994 HCAPLUS
(45) Evans, R; Science 1988, V240, P889 HCAPLUS
(46) Fawell, S; Cell 1990, V60, P953 HCAPLUS
(47) Forman, B; Mol Endocrinol 1990, V4, P1293 HCAPLUS
(48) Gewirth, D; Nature Structural Biology 1995, V2, P386 HCAPLUS
(49) Glass, C; Enocr Rev 1994, V15, P391 HCAPLUS
(50) Hajduk; Science 1997, V278, P497 HCAPLUS
(51) Hayashi, Y; J Clin Invest 1994, V94, P607 HCAPLUS
(52) Heery, E; Nature 1997, V387, P733
(53) Hollenberg; J Biol Chem 1995, V270(24), P14274 HCAPLUS
(54) Horwitz, K; Endocrine Rev 1992, V13, P146 HCAPLUS
(55) Jackson, R; Seminars in Oncology 1997, V24(2), PL164
(56) Jancarik; J Appl Crystallogr 1991, V24, P409 HCAPLUS
(57) Janknecht, R; Proc Natl Acad Sci USA 1991, V88, P8972 HCAPLUS
(58) Jones, T; ACTA Cryst 1991, V47, P110
(59) Jones, T; J Med Chem 1996, V39(4), P904 HCAPLUS
(60) Jorgensen, E; Hormonal Peptides and Proteins 1978, P107 HCAPLUS
(61) Jorgenson; Thyroid Hormones 1978, P150
(62) Jorgenson; Thyroid Research 1976, V378, P303
(63) Kabsch, W; Appl Crystallogr 1993, V26, P795 HCAPLUS
(64) Kabsch, W; Biopolymers 1983, V22, P2577 HCAPLUS
(65) Kakizawa, T; J Biol Chem 1997, V272(38), P23799 HCAPLUS
(66) Kediel, S; Mol Cell Biol 1994, V14, P287
(67) Laskowski, R; J Appl Crystallogr 1993, V26, P283 HCAPLUS
(68) Latham, K; J Biol Chem 1991, V256, P12088
(69) Laudet, V; Embo J 1992, V11, P1003 HCAPLUS
(70) Ledouarin, B; Embo J 1995, V14, P2020 HCAPLUS
(71) Lee, J; Molec Endocrinol 1995, V9, P243 HCAPLUS
(72) Lee, J; Nature 1995, V374, P91 HCAPLUS
(73) Leeson, P; J Med Chem 1987, V31, P37
(74) Leeson, P; J Med Chem 1989, V32, P320 HCAPLUS
(75) Leitman, D; J Biol Chem 1991, V266, P9343 HCAPLUS
(76) Leng, X; J Steroid Biochem Molec Biol 1993, V46, P643 HCAPLUS
(77) Leng, X; Mol and Cellular Biol 1995, V15, P255 HCAPLUS
(78) Lewis, N; Synthesis 1987, P1103 HCAPLUS
(79) Lin, B; Mol Cell Biol 1997, V17(10), P6131 HCAPLUS
(80) Lin, K; Mol Endocrinol 1991, V5, P485 HCAPLUS
(81) Lleywegt, G; ESF/CCP4 Newsletter 1994, P20
(82) Luisi, B; Nature 1991, V352, P497 HCAPLUS
(83) McGrath; Biotechniques 1989, V7, P246 HCAPLUS
(84) McGrath, M; J Mol Biol 1994, V237, P236 HCAPLUS
(85) McRee; Practical Protein Crystallography, Chapters 1, 2 and 3 1993
(86) Meier, C; Mol Endocrinol 1992, V6, P248 HCAPLUS
(87) Miura; US 5116828 1992 HCAPLUS
(88) Monaco; Science 1995, V268, P1039 HCAPLUS
(89) Murshudov; Refinement of Protein Structures, Proceedings of Daresbury Study Weekend 1996, P1
(90) Navaza, J; Acta Crytallographica Section A-Fundamentals of Crystallography 1994, V50, P157
(91) Nicholls, A; Proteins 1991, V11, P281 HCAPLUS
(92) Onate, S; Science V270, P1354 HCAPLUS
(93) Otwinoski, Z; Proceedings of the CCP4 Study Weekend 1991, P80
(94) Otwinoski, Z; Proceedings of the CCP4 Study Weekend:Data Collection and Processing 1993, P56
(95) Ozato; US 5403925 1995 HCAPLUS
(96) O'Donnell, A; Mol Endocrinol 1991, V5, P94 HCAPLUS
(97) Rastinejad, R; Nature 1995, V375, P203

(98) Raynaud, J; J Steroid Biochem 1986, V25, P811 HCAPLUS
(99) Refetoff, S; Endocr Rev 1993, V14, P348 HCAPLUS
(100) Ribeiro; Recent Progress in Hormone Research 1998, V53, P351 HCAPLUS
(101) Ribeiro, R; Ann N Y Acad Sci 1995, V758, P366 HCAPLUS
(102) Ribeiro, R; Annu Rev Med 1995, V46, P443 HCAPLUS
(103) Ribeiro, R; Mol Endocrinol 1992, V6, P1142 HCAPLUS
(104) Robsseau, G; J Mol Biol 1972, V67, P99
(105) Saatcioglu, F; Mol Cell Biol 1993, V13, P3675 MEDLINE
(106) Schwabe, J; Cell 1993, V75, P567 HCAPLUS
(107) Seielstad; Molecular Endocrinology 1995, V9, P647 HCAPLUS
(108) Selmi, S; J Biol Chem 1991, V266, P11589 HCAPLUS
(109) Shibata, H; Recent Progress in Hormone Res 1997, V52, P141 MEDLINE
(110) Stephan; Biochem Pharmacol 1992, V13, P1969
(111) Swaffield, J; Nature 1995, V374, P88 HCAPLUS
(112) Tagami, T; Mol Cell Biol 1997, V17(5), P2642 HCAPLUS
(113) Toney; J; Biochemistry 1993, V32, P2 HCAPLUS
(114) Tsai, M; Annu Rev Biochem 1994, V63, P451 HCAPLUS
(115) Wagner; Nature 1995, V378(6558), P670
(116) Westerfield, W; Endocrinology 1965, V77, P802
(117) Yokoyama; J Med Chem 1995, V38, P695 HCAPLUS
(118) Zenkie, M; Cell 1990, V61, P1035
(119) Zhu, X; J Biol Chem 1997, V272(14), P9048 HCAPLUS

L86 ANSWER 2 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 2001:284219 HCAPLUS
DN 134:305782
TI Crystallographic **structure** of the androgen receptor **ligand binding** domain and pharmacological applications
IN Weinmann, Roberto; Einspahr, Howard M.; Krystek, Stanley R., Jr.; Sack, John S.; Salvati, Mark E.; Tokarski, John S.; Wang, Chihuei; Attar, Ricardo M.
PA Bristol-Myers Squibb Company, USA
SO PCT Int. Appl., 83 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM G01N033-53
CC 2-2 (Mammalian Hormones)
Section cross-reference(s): 1, 75
FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2001027622	A1	20010419	WO 2000-US28495	20001013 <--
	W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
	RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
	EP 1222459	A1	20020717	EP 2000-972172	20001013 <--
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL				
	JP 2003511467	T2	20030325	JP 2001-530581	20001013 <--
PRAI	US 1999-159394P	P	19991014 <--		
	WO 2000-US28495	W	20001013		

AB The first crystal **structure** of the androgen receptor **ligand binding** domain has been detd. to 2.0 **angstrom** resoln. Disclosed are the **coordinates** for the crystal **structure**, and methods for detg. agonists, partial agonists, antagonists, partial antagonists, and selective androgen

receptors modulators of the androgen receptor.
ST androgen receptor **ligand** domain crystal **structure** pharmacol
IT Aging, animal
(age-related disease treatment; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT **Computer application**
Crystal growth
Crystal **structure**
Drug design
Drug screening
Mammal (Mammalia)
Molecular modeling
Molecular recognition
Protein sequences
Rat
(crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT Androgen receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
(crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT **Ligands**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT Sexual behavior
(impotence, treatment of; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT Prostate gland
(neoplasm, inhibitors; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT Antitumor agents
(prostate gland; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT **Conformation**
Helix (conformation)
(**protein**; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT Information systems
(storage; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT Osteoporosis
(therapeutic agents; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT 335182-97-3D, 672-917-Androgen receptor (rat), mutants are claimed
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
(**amino acid** sequence; crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT 521-18-6D, Dihydrotestosterone, complexes with androgen receptor **ligand binding** domain
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP

(Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
(crystallog. **structure** of androgen receptor **ligand binding** domain and pharmacol. applications)
IT 335190-37-9 335190-38-0
RL: PRP (Properties)
(unclaimed nucleotide sequence; crystallog. **structure** of the androgen receptor **ligand binding** domain and pharmacol. applications)
RE.CNT 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Claussner; US 5434176 A 1995 HCAPLUS
(2) Dedhar; US 5854202 A 1998 HCAPLUS
(3) Ekena; Journal of Biological Chemistry 1998, V273(2), P693 HCAPLUS
(4) Evans; US 5298429 A 1994 HCAPLUS
(5) Gaillard-Kelly; US 5411981 A 1995 HCAPLUS
(6) Jones; US 5693646 A 1997 HCAPLUS
(7) Miyamoto; Proceedings of the National Academy of Sciences 1998, V95, P7379 HCAPLUS
(8) Nakada; Cancer Research 1993, V53, P1967 HCAPLUS
(9) Sai; American Journal of Human Genetics 1990, V46, P1095 HCAPLUS

L86 ANSWER 3 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 2000:145057 HCAPLUS
DN 132:205136
TI Methods and systems for predicting **protein** function
IN Skolnick, Jeffrey; Fetrow, Jacquelyn S.
PA The Scripps Research Institute, USA
SO PCT Int. Appl., 173 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM C12Q001-00
ICS C12Q001-26; C12Q001-37; C12Q001-44; C12N009-00; C12N009-02; C12N009-14; C12N009-22; C12N009-48; C12N009-50; G01N031-00; G01N033-00; A61K038-00; C07K001-00
CC 9-16 (Biochemical Methods)
Section cross-reference(s): 3, 6, **7**
FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE	
PI	WO 2000011206	A1	20000302	WO 1999-US11913	19990527	<--
	W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM					
	RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG					
	CA 2340284	AA	20000302	CA 1999-2340284	19990527	<--
	AU 9942187	A1	20000314	AU 1999-42187	19990527	<--
	EP 1108055	A1	20010620	EP 1999-926014	19990527	<--
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI					
	JP 2002523057	T2	20020730	JP 2000-566458	19990527	<--
	US 2001034580	A1	20011025	US 2001-839821	20010420	<--
PRAI	US 1998-99300P	P	19980825	<--		
	US 1999-120311P	P	19990216	<--		
	US 1999-322067	A3	19990527	<--		
	WO 1999-US11913	W	19990527	<--		

AB The present invention concerns methods and systems for predicting the

biol. function(s) of **proteins**. The invention is based on the development of functional site descriptors for discrete **protein** biol. functions. Functional site descriptors are geometric representations of **protein** functional sites in three-dimensional space, and can also include addnl. parameters, for example, **conformational** information. Following their development, one or more functional site descriptors (for one or more different biol. functions) are used to probe **protein structures** to det. if such **structures** contain the functional sites described by the corresponding functional site descriptors. If so, the **protein**(s) contg. the functional site(s) are predicted to have the corresponding biol. function(s). In preferred embodiments, a library of functional site descriptors is used to probe inexact **protein structures** derived by **computational** methods from **amino acid** sequence information to predict the biol. function(s) of such sequences and of the gene(s) encoding the same. The Escherichia coli genome was **screened** by the sequence-to-**structure**-to-function paradigm to identify **proteins** having disulfide oxidoreductase activity.

ST **protein** function prediction system; sequence library functional site descriptor **protein**; **computer** program **protein** function prediction; genome sequence **screening** disulfide oxidoreductase **protein**

IT Mycoplasma genitalium
(**algorithm** in functional anal. of; methods and systems for predicting **protein** function)

IT **Ligands**
RL: BAC (Biological activity or effector, except adverse); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(**binding** domain for, as functional site descriptor; methods and systems for predicting **protein** function)

IT **Enzyme** functional sites
(cofactor-**binding**, domain for, as functional site descriptor; methods and systems for predicting **protein** function)

IT Thioredoxins
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(disulfide oxidoreductase activity of, functional site descriptors of; methods and systems for predicting **protein** function)

IT **Molecular association**
(domain for, as functional site descriptor; methods and systems for predicting **protein** function)

IT Haemophilus influenzae
Methanococcus jannaschii
(functional anal. of genome of, for **proteins** having thiol/disulfide oxidoreductase activity; methods and systems for predicting **protein** function)

IT Escherichia coli
(functional **screening** of genome of, for disulfide oxidoreductase activity; methods and systems for predicting **protein** function)

IT Genome
(functional **screening** of, of Escherichia coli, for disulfide oxidoreductase activity; methods and systems for predicting **protein** function)

IT Antigens
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
(functional site descriptor for **protein** domain **binding** to; methods and systems for predicting **protein** function)

IT **Proteins**, specific or class

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(glutaredoxins, disulfide oxidoreductase activity of, functional site descriptors of; methods and systems for predicting **protein** function)

IT **Algorithm**
Computer program
Computers
DNA sequences
Enzyme functional sites
NMR (nuclear magnetic resonance)
Protein folding
Protein motifs
Protein sequences
Simulation and **Modeling**, biological
Structure-activity relationship
(methods and systems for predicting **protein** function)

IT **Proteins**, general, properties
RL: PRP (Properties)
(methods and systems for predicting **protein** function)

IT **Peptide** library
(of functional site descriptors; methods and systems for predicting **protein** function)

IT Gene
RL: PRP (Properties)
(open reading frame, functional **screening** of, of Escherichia coli, for disulfide oxidoreductase activity; methods and systems for predicting **protein** function)

IT Animal
Canidae
Cattle
Felidae
Horse (Equus caballus)
Mammal (Mammalia)
Plant (Embryophyta)
Prokaryote
Sheep
Swine
Virus
(**protein** of; methods and systems for predicting **protein** function)

IT **Secondary structure**
Tertiary structure
(**protein**; methods and systems for predicting **protein** function)

IT Saccharomyces cerevisiae
(**proteins** of, disulfide oxidoreductase functional site descriptor testing against; methods and systems for predicting **protein** function)

IT Toxins
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(ribo-, T1 RNase functional site descriptor testing against; methods and systems for predicting **protein** function)

IT Information systems
(searching; methods and systems for predicting **protein** function)

IT **Enzyme** functional sites
(substrate-**binding**, domain for, as functional site descriptor; methods and systems for predicting **protein** function)

IT Crystallography
(x-ray; methods and systems for predicting **protein** function)

IT 9013-93-8, Phospholipase 9026-12-4 9068-62-6, Disulfide reductase
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(functional site descriptor for; methods and systems for predicting **protein** function)
IT 52-90-4, Cysteine, properties 56-40-6, Glycine, properties 56-41-7, L-Alanine, properties 56-45-1, Serine, properties 56-84-8, Aspartic acid, properties 56-85-9, Glutamine, properties 56-86-0, L-Glutamic acid, properties 56-87-1, L-Lysine, properties 60-18-4, Tyrosine, properties 61-90-5, Leucine, properties 63-68-3, Methionine, properties 63-91-2, Phenylalanine, properties 70-47-3, Asparagine, properties 71-00-1, Histidine, properties 72-18-4, Valine, properties 72-19-5, Threonine, properties 73-22-3, Tryptophan, properties 73-32-5, Isoleucine, properties 74-79-3, Arginine, properties 147-85-3, Proline, properties
RL: PRP (Properties)
(in functional site descriptor; methods and systems for predicting **protein** function)
IT 9001-62-1, Lipase
RL: PRP (Properties)
(**structure** of, in building consensus active site **structure** for .alpha./.beta. hydrolases; methods and systems for predicting **protein** function)
IT 94587-93-6 117698-18-7 132052-53-0, .alpha.-Sarcin (Aspergillus giganteus strain MDH18894 precursor reduced) 136073-15-9, Restrictocin (Aspergillus restrictus clone pFB39c1 precursor reduced) 172142-71-1 180256-67-1 260345-59-3 260345-60-6 260345-61-7 260345-62-8 260345-63-9 260345-64-0
RL: PRP (Properties)
(unclaimed sequence; methods and systems for predicting **protein** function)
IT 9027-41-2, Hydrolase
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(.alpha./.beta.-**fold** contg., functional site descriptor for; methods and systems for predicting **protein** function)
RE.CNT 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Biswanadhan; Biochemistry 1991, V30, P11164
(2) Chao; CABIOS 1992, V8(5), P481 HCAPLUS
(3) Clarke; Science 1988, V240, P521 HCAPLUS
(4) Escalier; Journal of Computational Biology 1998, V5(1), P41 HCAPLUS
(5) Kalhammer; FEBS Letters 1997, V414, P599 HCAPLUS
(6) Pawlowski; The Heliobacter pylori genome:From sequence analysis to structural and functional predictions Proteins: Structure, Function and Genetics 1999, V36, P20 HCAPLUS
(7) Rychlewski; Functional insight from structural predictionS: Analysis of the Escherichia coli genome 1999, V8, P614 HCAPLUS
(8) Wallace; Protein Science 1996, V5, P1001 HCAPLUS
(9) Zhang; Protein Science 1999, V8, P1104 HCAPLUS

L86 ANSWER 4 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **2000:68632** HCAPLUS
DN **132:105024**
TI Determination of **ligands** for **proteins** via **modeling** of **molecular surface patches**
IN **Frommel, Cornelius; Preissner, Robert; Goede, Andrean**
PA **Jerini Bio Tools G.m.b.H., Germany**
SO PCT Int. Appl., 21 pp.
CODEN: PIXXD2
DT Patent
LA German

IC ICM G01N033-53
ICS C07K001-00
CC 9-16 (Biochemical Methods)
Section cross-reference(s): 1, 6
FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2000004380	A1	20000127	WO 1999-EP4951	19990713 <--
	W: JP, US				
	RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
	DE 19831758	A1	20000203	DE 1998-19831758	19980715 <--
	EP 1095272	A1	20010502	EP 1999-934689	19990713 <--
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
	US 2002048776	A1	20020425	US 2001-772538	20010129 <--
PRAI	DE 1998-19831758	A	19980715 <--		
	WO 1999-EP4951	W	19990713 <--		

AB The invention relates to a method for detg. **ligands** for **proteins**. Said method comprises detg., by means of **secondary structural** elements of a given **protein** which form the **binding** site, **mol. surface patches** (MSP) which are compared with known **mol. surface patches** with **ligand**. The method is used for drug **design**, in biotechnol. for finding affinity purifn. **ligands**, in the food industry for the research of flavoring compds. in relation to taste receptors.

ST **protein ligand binding** drug **design** affinity purifn flavoring material

IT Affinity chromatography
Computer application
Databases
Drug design
Flavoring materials
Molecular association
Molecular recognition
Molecular surface
Protein motifs
Simulation and Modeling, physicochemical
(detn. of **ligands** for **proteins** via **modeling** of **mol. surface patches**)

IT Taste receptors
RL: FFD (Food or feed use); BIOL (Biological study); USES (Uses)
(detn. of **ligands** for **proteins** via **modeling** of **mol. surface patches**)

IT **Enzymes**, properties
RL: PRP (Properties)
(detn. of **ligands** for **proteins** via **modeling** of **mol. surface patches**)

IT **Ligands**
RL: PRP (Properties)
(detn. of **ligands** for **proteins** via **modeling** of **mol. surface patches**)

IT **Proteins**, general, properties
RL: PRP (Properties)
(detn. of **ligands** for **proteins** via **modeling** of **mol. surface patches**)

IT **Peptides**, properties
RL: PRP (Properties)
(**ligands**; detn. of **ligands** for **proteins** via **modeling** of **mol. surface patches**)

IT **Conformation**

Secondary structure
(**protein**; detn. of **ligands** for **proteins**
via **modeling** of **mol. surface**
patches)

RE.CNT 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Creative Biomolecules Inc; WO 9726277 A 1997 HCAPLUS
(2) Delisi, C; US 5495423 A 1996
(3) Goede; JOURNAL OF COMPUTATIONAL CHEMISTRY 1997, V18(9), P1113 HCAPLUS
(4) Preissner; JOURNAL OF MOLECULAR BIOLOGY 1998, V280(3), P535 MEDLINE
(5) Preissner, R; FEBS LETTERS 1997, V414(2), P425 HCAPLUS
(6) Scripps Clinic Res; WO 9116683 A 1991 HCAPLUS
(7) Scripps Research Inst; WO 9321206 A 1993 HCAPLUS
(8) Srinivasan, S; US 5557535 A 1996 HCAPLUS
(9) Univ California; WO 9301484 A 1993 HCAPLUS

L86 ANSWER 5 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 2000:53689 HCAPLUS
DN 132:106951
TI Isolated amphiphilic **peptides** derived from the cytoplasmic tail of viral envelope **proteins**
IN Rozenberg, Yanina; Anderson, W. French
PA University of Southern California, USA
SO PCT Int. Appl., 83 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM C07K014-00
CC 15-2 (Immunochemistry)
Section cross-reference(s): 3, 63
FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2000002909	A2	20000120	WO 1999-IB1261	19990708 <--
	WO 2000002909	A3	20000504		
	W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
	RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
	CA 2333855	AA	20000120	CA 1999-2333855	19990708 <--
	AU 9943869	A1	20000201	AU 1999-43869	19990708 <--
	EP 1093519	A2	20010425	EP 1999-926703	19990708 <--
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
	US 2002058020	A1	20020516	US 2001-756250	20010108 <--
PRAI	US 1998-112544	A	19980709 <--		
	WO 1999-IB1261	W	19990708 <--		

AB An isolated **peptide** comprising an **amino acid** sequence derived from a viral envelope **protein**, wherein at least a portion of the **amino acid** sequence is located within the cytoplasmic tail or membrane-spanning region of a viral envelope **protein**. Such **peptides** are amphiphilic in nature, provide for the destabilization of membranes, and facilitate the entry of viral particles into cells and the efficient formation of viral particles. The **peptides** may, in another embodiment, be attached to the viral membrane, along with a targeting **polypeptide**, as part of an artificial viral envelope **protein**.
ST virus retrovirus cytoplasmic envelope **peptide** liposome

IT **Algorithm**
(Chou, Fasman and Rose; amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT **Algorithm**
(Kyte and Doolittle; amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT Biomarkers (biological responses)
Circular dichroism
DNA sequences
Drug targeting
ESR (electron spin resonance)
Gene targeting
Liposomes
Murine leukemia virus
Protein sequences
Retroviral vectors
Retroviridae
Virus
Virus vectors
.alpha.-Helix
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT Polynucleotides
RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT **Ligands**
gag **proteins**
RL: BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT Envelope **proteins**
RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT **Peptides**, biological studies
RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT **Enzymes**, biological studies
RL: BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(gene pol; amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT Drug delivery systems
(liposomes; amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT Animal cell
Animal cell line
(pre-packaging; amphiphilic **peptides** of viral envelope

proteins and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT Membrane, biological
(viral; amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT 255719-95-0, Ilnrlvqfvkdrisvvqal **peptide**+ 255719-96-1, Lkvlttglpalms **peptide**+ 255719-97-2, Lkvlttglpalmswi **peptide**+
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT 255859-47-3 255859-48-4
RL: BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT 26853-31-6, POPC 81490-05-3, POPG
RL: BUU (Biological use, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

IT 255859-93-9
RL: PRP (Properties)
(nucleotide sequence; amphiphilic **peptides** of viral envelope **proteins** and encoding polynucleotides for anal. and targeting therapeutic **peptides** and gene)

L86 ANSWER 6 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 2000:5177 HCAPLUS
DN 132:136203
TI Fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-**ligand binding**
AU Burrows, Louise; Clark, Katherine; Mould, A. Paul; Humphries, Martin J.
CS Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, Manchester, M13 9PT, UK
SO Biochemical Journal (**1999**), 344(2), 527-533
CODEN: BIJOAK; ISSN: 0264-6021
PB Portland Press Ltd.
DT Journal
LA English
CC 15-3 (Immunochemistry)
Section cross-reference(s): 6
AB The high-affinity interaction of integrin .alpha.5.beta.1 with the central cell-**binding** domain of fibronectin requires both the Arg-Gly-Asp (RGD) sequence (in the tenth type III repeat) and a **second** site Pro-His-Ser-Arg-Asn (PHSRN) in the adjacent ninth type III repeat, which synergizes with RGD. Arg-Arg-Glu-Thr-Ala-Trp-Ala (RRETAWA) is a novel peptidic **ligand** for .alpha.5.beta.1, identified by phage display, which blocks .alpha.5.beta.1-mediated cell adhesion to fibronectin. A key question is the location of the **binding** sites for these **ligand** sequences within the integrin. In this study we have identified residues that form part of the epitopes of three inhibitory anti-.alpha.5 monoclonal antibodies (mAbs): 16, P1D6 and SNAKA52. These mAbs have distinct functional properties. MAb 16 blocks the **recognition** of RGD and RRETAWA, whereas P1D6 blocks **binding** to the synergy sequence. The **binding** of SNAKA52 is inhibited by anti-.beta.1 mAbs, indicating that its epitope is close to

the interface between the .alpha. and .beta. subunits. Residues in human .alpha.5 were replaced with the corresponding residues in mouse .alpha.5 by site-directed mutagenesis; wild-type or mutant human .alpha.5 was expressed on the **surface** of .alpha.5-deficient Chinese hamster ovary cells. MAb **binding** was assessed by flow cytometry and by adhesion to the central cell-**binding** domain of fibronectin or RRETAWA by cell attachment assay. All three epitopes were located to different putative loops in the N-terminal domain of .alpha.5. As expected, disruption of these epitopes had no effect on **ligand recognition** by .alpha.5.beta.1. The locations of these epitopes are consistent with the .beta.-propeller **model** for integrin .alpha.-subunit **structure** and allow us to propose a topol. image of the integrin-**ligand** complex.

ST epitope mapping monoclonal Ig integrin **ligand binding**; fibronectin integrin **binding conformation** propeller **model**

IT Animal cell line
(CHO; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)

IT Immunoglobulins
RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process)
(G, monoclonal; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)

IT Extracellular matrix
Molecular association
Molecular topology
Repeat motifs (protein)
(fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)

IT **Peptides**, biological studies
RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process)
(fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)

IT Fibrinogens
Ligands
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)

IT **Secondary structure**
(fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding** and human and mouse blades 2 and 3 of integrin .beta.-propeller **model**)

IT **Molecular modeling**
(fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding** and human and mouse blades 2 and 3 of .beta.-propeller **model**)

IT **Protein** sequences
(homol.; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding** and human and mouse blades 2 and 3 of .beta.-propeller **model**)

IT Epitopes
(mapping; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)
IT **Conformation**
(**protein**, .beta.-propeller **model**; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding** and human and mouse blades 2 and 3 of integrin .beta.-propeller **model**)
IT Mutagenesis
(site-directed; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding** studied by)
IT Integrins
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(.alpha.5.beta.1; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)
IT **Conformation**
(.beta.-strand, seven bladed .beta.-propeller; fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)
IT 99896-85-2P 158622-13-0P 168179-93-9P
RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process)
(fine mapping of inhibitory anti-.alpha.5 monoclonal antibody epitopes that differentially affect integrin-fibronectin **ligand binding**)
RE.CNT 32 THERE ARE 32 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Aota, S; J Biol Chem 1994, V269, P24756 HCAPLUS
(2) Copie, V; J Mol Biol 1998, V277, P663 HCAPLUS
(3) Danen, E; J Biol Chem 1995, V270, P21612 HCAPLUS
(4) Higgins, D; Comput Appl Biosci 1991, V8, P189
(5) Humphries, M; Trends Cell Biol 1998, V8, P78 HCAPLUS
(6) Irie, A; Proc Natl Acad Sci USA 1997, V94, P7198 HCAPLUS
(7) Kamata, T; Biochem J 1995, V305, P945 HCAPLUS
(8) Kamata, T; J Biol Chem 1996, V271, P18610 HCAPLUS
(9) Koivunen, E; J Cell Biol 1994, V124, P373 HCAPLUS
(10) Krukonis, E; J Biol Chem 1998, V273, P31837 HCAPLUS
(11) Kunkel, T; Methods Enzymol 1987, V154, P367 HCAPLUS
(12) Leahy, D; Cell 1996, V84, P155 HCAPLUS
(13) Lee, J; Cell 1995, V80, P631 HCAPLUS
(14) Loftus, J; J Biol Chem 1996, V271, P2033 HCAPLUS
(15) Loftus, J; J Clin Invest 1997, V99, P2302 HCAPLUS
(16) Mould, A; Biochem J 1998, V331, P821 HCAPLUS
(17) Mould, A; J Biol Chem 1995, V270, P26270 HCAPLUS
(18) Mould, A; J Biol Chem 1996, V271, P20365 HCAPLUS
(19) Mould, A; J Biol Chem 1997, V272, P17283 HCAPLUS
(20) Mould, A; J Biol Chem 1998, V273, P25664 HCAPLUS
(21) Newham, P; Mol Med Today 1996, V2, P304 HCAPLUS
(22) Obara, M; Cell 1988, V53, P649 HCAPLUS
(23) Pierschbacher, M; Nature (London) 1984, V309, P30 HCAPLUS
(24) Schiffer, S; J Biol Chem 1995, V270, P14270 HCAPLUS
(25) Schreiner, C; J Cell Biol 1989, V109, P3157 HCAPLUS
(26) Springer, T; Proc Natl Acad Sci USA 1997, V94, P65 HCAPLUS
(27) Takada, Y; Matrix Biol 1997, V16, P143 HCAPLUS
(28) Tozer, E; J Biol Chem 1996, V271, P21978 HCAPLUS
(29) Tuckwell, D; Biochem J 1992, V285, P325 HCAPLUS

(30) Tuckwell, D; Cell Adhes Commun 1994, V2, P385 HCAPLUS
(31) Tuckwell, D; FEBS Lett 1997, V400, P297 HCAPLUS
(32) Yamada, K; J Cell Biol 1984, V99, P29 HCAPLUS

L86 ANSWER 7 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1999:641074 HCAPLUS
DN 131:282013
TI Methods and compounds for modulating nuclear receptor activity
IN Shiau, Andrew; Kushner, Peter J.; Agard, David A.; Greene, Geoffrey L.
PA University of California, USA; Arch Development Corp.
SO PCT Int. Appl., 207 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM G01N033-48
CC 2-1 (Mammalian Hormones)
Section cross-reference(s): 1
FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 9950658	A2	19991007	WO 1999-US6937	19990330 <--
	W: AU, CA, JP, KR, US				
	RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
	CA 2324060	AA	19991007	CA 1999-2324060	19990330 <--
	AU 9934571	A1	19991018	AU 1999-34571	19990330 <--
	CA 2323575	AA	19991125	CA 1999-2323575	19990330 <--
	WO 9960014	A2	19991125	WO 1999-US6899	19990330 <--
	WO 9960014	A3	20000518		
	WO 9960014	C2	20020829		
	W: AU, CA, JP, KR				
	RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
	EP 1068529	A2	20010117	EP 1999-944980	19990330 <--
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
	EP 1144997	A2	20011017	EP 1999-916206	19990330 <--
	EP 1144997	A3	20020828		
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
	US 2002061539	A1	20020523	US 1999-281717	19990330 <--
	JP 2002516983	T2	20020611	JP 2000-541516	19990330 <--
PRAI	US 1998-79956P	P	19980330 <--		
	US 1998-113014P	P	19981216 <--		
	US 1998-113146P	P	19981216 <--		
	US 1998-79965P	P	19980330 <--		
	WO 1999-US6899	W	19990330 <--		
	WO 1999-US6937	W	19990330 <--		

AB The present invention relates to methods and agonist/antagonist compds. for modulating nuclear receptor activity, and nuclear receptor **ligand binding**. The invention includes a method for identifying residues comprising a **ligand binding** domain for a nuclear receptor of interest. Also included in a method of identifying agonists and/or antagonists that **bind** to the **ligand binding** domain of the nuclear receptors, and the estrogen receptor in particular. The invention is exemplified by identification and manipulation of the **ligand binding** domain of the estrogen receptor and compds. that **bind** to this site. The methods can be applied to other nuclear receptors including TR, GR and PR.

ST estrogen nuclear receptor **ligand screening**
IT **Bond**
(hydrophobic; **screening** for compds. modulating nuclear

receptor activity)
IT **Helix (conformation)**
(**protein; screening** for compds. modulating nuclear receptor activity)
IT Peroxisome
(receptors; **screening** for compds. modulating nuclear receptor activity)
IT Combinatorial library
Computer application
Crystal **structure**
Dipole
Drug screening
Hydrogen bond
Molecular association
Molecular cloning
Molecular modeling
Peptidomimetics
Protein sequences
(**screening** for compds. modulating nuclear receptor activity)
IT Androgen receptors
Estrogen receptors
Glucocorticoid receptors
Ligands
Mineralocorticoid receptors
Nuclear receptors
Peptides, biological studies
Progesterone receptors
Retinoid receptors
Thyroid hormone receptors
Vitamin D receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(**screening** for compds. modulating nuclear receptor activity)
IT Information systems
(storage; **screening** for compds. modulating nuclear receptor activity)
IT **Bond**
(van der Waals; **screening** for compds. modulating nuclear receptor activity)
IT 61-90-5, Leucine, biological studies 63-68-3, Methionine, biological studies 71-00-1, Histidine, biological studies
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); BIOL (Biological study); OCCU (Occurrence)
(of human estrogen receptor .alpha.; **screening** for compds. modulating nuclear receptor activity)
IT 50-28-2, 17.beta.-Estradiol, biological studies 56-53-1, Diethylstilbestrol 84-16-2, Mesohexestrol 143-50-0, Kepone 479-13-0, Coumestrol 789-02-6 1972-08-3, .DELTA.9-Thc 10540-29-1, Tamoxifen 17924-92-4, Zearalenone 34816-55-2, Moxestrol 84449-90-1, Raloxifene 98007-99-9 155701-61-4, Gw5638 182167-03-9, Em800 205128-72-9 245122-98-9 245122-99-0 245123-00-6 245123-01-7 245123-02-8 245123-03-9 245123-04-0 245123-05-1 245123-06-2 245123-07-3 245123-08-4 245123-09-5 245676-26-0 245676-32-8 245676-38-4 245676-43-1 245676-45-3 245676-46-4 245676-47-5 245676-49-7 245676-54-4 246236-26-0
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(**screening** for compds. modulating nuclear receptor activity)
IT 245742-95-4, PN: WO9950658 SEQID: 1 unclaimed **protein**
245742-96-5, PN: WO9950658 SEQID: 2 unclaimed **protein**
RL: PRP (Properties)
(unclaimed **protein** sequence; methods and compds. for modulating nuclear receptor activity)

L86 ANSWER 8 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1999:353941 HCAPLUS
DN 131:181313
TI **Protein surface recognition**
AU Salvatella, X.; Haack, T.; Gairi, M.; De Mendoza, J.; Peczuh, M. W.; Hamilton, A. D.; Giralt, E.
CS Departament de Quimica Organica, Universitat de Barcelona, Barcelona, 08028, Spain
SO NATO ASI Series, Series C: Mathematical and Physical Sciences (**1999**), 526(NMR in Supramolecular Chemistry), 267-280
CODEN: NSCSDW; ISSN: 0258-2023
PB Kluwer Academic Publishers
DT Journal
LA English
CC 6-3 (General Biochemistry)
AB The development of a **ligand** for **mol. recognition** of the external **protein surface** is a challenging task. Difficulties arise because of the high degree of solvation and the hydrophilic character of **protein surfaces**, where intermol. interactions such as hydrogen bonds and electrostatic forces play a key role. A tetraguanidinium compd. was studied for its ability to be used as a receptor for **recognition** of **peptides** or **proteins** with anionic helical **structures**. The **recognition** depends primarily on two factors, helicity of the **protein** and accessibility of neg. charged residues (aspartate or glutamate) at crit. positions to interact with the **ligand**. A search of the Brookhaven **Protein** Data Bank yielded 31 entries that contained the relevant features. Visual inspection of these **protein structures** narrowed the possible hits to 10, which were considered good candidates for **recognition** and were studied using a more systematic approach. For most of these **proteins**, the accessibility and helicity of the anionic paths were too low to expect any **binding** to the tetraguanidinic receptor. One exception was the p53 tumor suppressor **protein**, which was chosen to assess the **surface-binding** properties of the tetraguanidinic receptor.
ST **protein surface recognition** tetraguanidinium receptor
IT Annexins
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(IV; **protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT Carboxyl group
(accessibility; **protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT **Molecular recognition**
.alpha.-Helix
(**protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT **Proteins**, general, biological studies
p53 (**protein**)
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(**protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT Myosins
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(regulatory domain; **protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT Transcription factors

RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(repressors, of primer of DNA replication; **protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT 239136-49-3
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(p53 tumor suppressor **protein model peptide** ; **protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT 9040-57-7, Ribonucleotide reductase
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(**protein** subunit R2; **protein surface recognition** by a tetraguanidinium receptor **ligand**)
IT 9001-04-1, Pyruvate decarboxylase 9028-56-2, 3.alpha.-Hydroxysteroid dehydrogenase 9031-26-9, Lysyl-tRNA synthetase 37329-68-3, Procarboxypeptidase B 118901-79-4, Scytalone dehydratase 194547-39-2
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(**protein surface recognition** by a tetraguanidinium receptor **ligand**)
RE.CNT 17 THERE ARE 17 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Bennett, M; Biochemistry 1996, V35, P10702 HCAPLUS
(2) Clubb, R; Protein Sci 1995, V4, P855 HCAPLUS
(3) Connolly, M; Science 1983, V221, P709 HCAPLUS
(4) Dyda, F; Biochemistry 1993, V32, P6165 HCAPLUS
(5) Eberle, W; biochemistry 1990, V29, P7402 HCAPLUS
(6) Echavarren, A; J Am Chem Soc 1989, V111, P4994 HCAPLUS
(7) Hale, A; Eur J Biochem 1996, V236, P1 HCAPLUS
(8) Lee, W; Nature Struct Biol 1994, V1, P877 HCAPLUS
(9) Lundqvist, T; Structure 1994, V2, P937 HCAPLUS
(10) Nordlund, P; Nature 1990, V345, P593 HCAPLUS
(11) Onesti, S; J Mol Biol 1994, V243, P123 HCAPLUS
(12) Peczuh, M; J Am Chem Soc 1997, V119, P9327 HCAPLUS
(13) Sanchez-Quezada, J; J Am Chem Soc 1996, V118, P277
(14) Silverman, R; Mechanism-Based Enzyme Inactivation: Chemistry and Enzymology 1988
(15) Sutton, R; unpublished data
(16) Vendrell, J; Biochemistry 1990, V29, P7515 MEDLINE
(17) Xie, X; Nature 1994, V368, P306 HCAPLUS

L86 ANSWER 9 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1999:193896 HCAPLUS
DN 130:207008
TI Method and system for **protein modeling**
IN Srinivasan, Subhashini; Sudarsanam, Padmanaban
PA Immunex Corporation, USA
SO U.S., 29 pp., Cont.-in-part of U.S. 5,453,937.
CODEN: USXXAM
DT Patent
LA English
IC ICM G06F019-00
ICS C07K005-00; C07K014-00
NCL 702022000
CC 9-16 (Biochemical Methods)
Section cross-reference(s): 6
FAN.CNT 3

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5884230	A	19990316	US 1996-663809	19960614 <--

US 5453937 A 19950926 US 1993-55050 19930428 <--
US 5557535 A 19960917 US 1994-234812 19940428 <--
PRAI US 1993-55050 19930428 <--
US 1994-234812 19940428 <--
AB A method in a **computer** system for **modeling** a three-dimensional **structure** of a **model protein** is provided. In one embodiment, the **modeling** is based upon a three-dimensional **structure** of a template **protein** and an **amino acid** sequence alignment of the **model protein** and the template **protein**. For each **amino acid** in the **model protein**, when the template **protein** has an **amino acid** aligned with the **amino acid** of the **model protein**, the position of the backbone atom of the **amino acid** of the **model protein** is established based on the position of a topol. equiv. backbone atom in the aligned **amino acid** of the template **protein**. In another embodiment, the **modeling** of a variable region of the **model protein** is based on a collection of .psi. and .phi. **angle** values for **amino acid** pairs in a family of **proteins**. In a further embodiment, these .psi. and .phi. **angle** values are classified according to a tetramer of adjacent **amino acids** and filtered based on a most probable **conformation** of portions of the variable region of the **model protein**.
ST **protein** three dimensional **structure modeling**
IT **Algorithm**
Simulation and **Modeling**, physicochemical
(method and system for **protein modeling**)
IT **Amino acids**, properties
Proteins, general, properties
RL: PRP (Properties)
(method and system for **protein modeling**)
IT **Conformation**
(**protein**; method and system for **protein modeling**)
RE.CNT 18 THERE ARE 18 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Baldi; Proceeding of the National Academy of Science 1994, V91, P1059 HCAPLUS
(2) Blanco; US 5260882 1993
(3) Brooks; Journal of Computational Chemistry 1983, V4(2), P187 HCAPLUS
(4) Brooks, B; Journal of Computational Chemistry 1983, V4(2), P187 HCAPLUS
(5) Cramer; US 5025388 1991 HCAPLUS
(6) Cramer; US 5307287 1994 HCAPLUS
(7) Eisenberg; US 5436850 1995 HCAPLUS
(8) Havel, T; J Mol Biol 1991, P1 HCAPLUS
(9) Ladner; US 4881175 1989 HCAPLUS
(10) Lee; US 5241470 1993 HCAPLUS
(11) Moult, J; James Proteins:Structure, Function and Genetics 1986, V1, P146 HCAPLUS
(12) Pantoliano; US 4908773 1990 HCAPLUS
(13) Soukhanov, A; New Riverside University Dictionary 1988
(14) Srinivasan; US 5557535 1996 HCAPLUS
(15) Srinivasan; Protein Engineering 1993, V6(5), P501 HCAPLUS
(16) Srinivasan, S; International Journal of Quantum Chemistry:Quantum Biology Symposium 1986, V13, P167 HCAPLUS
(17) Still; US 5420805 1995 HCAPLUS
(18) Taylor, W; J Mol Biol 1986, V188, P233 HCAPLUS

L86 ANSWER 10 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1999:83909** HCAPLUS
DN **130:237856**

TI **Density** functional calculation of the electronic **structure** for **peptide-ligand** interactions
AU Bohr, H.; Jalkanen, Karl
CS Institute of Physics, b.307, The Technical University od Denmark, lyngby, 2800, Den.
SO Condensed Matter Theories (**1998**), 13, 95-112
CODEN: CMTHEO; ISSN: 0893-861X
PB Nova Science Publishers, Inc.
DT Journal
LA English
CC 34-3 (**Amino Acids, Peptides,** and **Proteins**)
Section cross-reference(s): 22
AB Quantum mech. calcns. of the electronic **structure** of di- and **tripeptides** in aq. soln. were carried out using d. functional techniques. The **conformational** energies, bond lengths, dihedral **angles,** and spectroscopical properties of zwitterionic **structures** stabilized by different amts. of H2O **mols.** were listed for L-alanyl-L-alanine. The stable zwitterionic **structures** given here were quite different from those reported by Barron et al. (1991). The usefulness of neural networks for the prediction of **secondary structures** was demonstrated for N-acetyl-L-alanine N'-methylamide.
ST **peptide** aq soln electronic **structure density** functional theory; **conformational** energy **peptide** aq soln DFT
IT **Conformational** potential
Density functional theory
Electronic **structure**
(electronic **structure** and **conformational** energy of di- and **tripeptides** in aq. soln. calcd. by d. functional techniques)
IT **Peptides,** properties
RL: PRP (Properties)
(electronic **structure** and **conformational** energy of di- and **tripeptides** in aq. soln. calcd. by d. functional techniques)
IT **Simulation and Modeling, physicochemical**
(neural network; electronic **structure** of di- and **tripeptides** in aq. soln. calcd. by d. functional techniques using neural networks)
IT Circular dichroism
Molecular structure
Vibrational spectra
(of di- and **tripeptides** in aq. soln. calcd. by d. functional techniques)
IT 56-41-7, L-Alanine, properties 1948-31-8 19701-83-8, N-Acetyl-L-alanine N'-methylamide
RL: PRP (Properties)
(electronic **structure** and **conformational** energy of di- and **tripeptides** in aq. soln. calcd. by d. functional techniques)
RE.CNT 54 THERE ARE 54 CITED REFERENCES AVAILABLE FOR THIS RECORD
RE
(1) Amos, R; Chem Phys L 1987, V133, P21 HCAPLUS
(2) Amos, R; J Phys Chem 1988, V92, P5571 HCAPLUS
(3) Barron, L; Spectrochimica Acta 1991, V47A, P1001 HCAPLUS
(4) Becke, A; J Chem Phys 1992, V96, P2155 HCAPLUS
(5) Becke, A; J Chem Phys 1992, V97, P9173 HCAPLUS
(6) Becke, A; J Chem Phys 1993, V98, P1372 HCAPLUS
(7) Becke, A; J Chem Phys 1993, V98, P5648 HCAPLUS
(8) Becke, A; Phys Rev A 1988, V38, P3098 HCAPLUS
(9) Bohr, H; FEBS Lett 1988, V241, P223 HCAPLUS

(10) Bohr, H; FEBS Lett 1990, V261, P43 HCAPLUS
(11) Buckingham, A; Chem Phys 1987, V112, P1 HCAPLUS
(12) Colwell, S; Chem Phys L 1993, V210, P261 HCAPLUS
(13) Colwell, S; Chem Phys L 1994, V217, P271
(14) Craw, J; J Phys Chem V100, P6304 HCAPLUS
(15) Diem, M; J Am Chem Soc 1988, V110, P6967 HCAPLUS
(16) Fortunelli, A; Chem Phys L 1994, V231, P34 HCAPLUS
(17) Freedman, T; J Am Chem Soc 1988, V110, P6970 HCAPLUS
(18) Godby, R; Phys Rev B 1988, V37, P10159
(19) Grabec, I; J Acoust Soc Am 1989, V85, P1226
(20) Hohenberg, P; Phys Rev A 1964, V136, P864
(21) Holley, L; Proc Natl Acad Sci USA 1989, V86, P152 HCAPLUS
(22) Jalkanen, K; "Density Functional and Neural Network Analysis of Hydration Effects and Spectroscopic and Structural Correlations in Small Peptides and Amino Acids" (To appear in "Computational Methods in Genome Research" 1996
(23) Jalkanen, K; Chem Phys in press
(24) Johnson, B; J Chem Phys 1994, V100, P7429 HCAPLUS
(25) Kohn, W; Phys Rev A 1965, V140, P1133
(26) Komornicki, A; J Chem Phys 1993, V98, P1398 HCAPLUS
(27) Lannoo, M; Phys Rev B 1985, V32, P3890 HCAPLUS
(28) Lee, A; Chem Phys L 1994, V229, P225 HCAPLUS
(29) Lee, C; J Chem Phys 1994, V100, P9018 HCAPLUS
(30) Lee, C; Phys Rev B 1988, V37, P785 HCAPLUS
(31) Levine, Z; Phys Rev L 1989, V63, P1719 HCAPLUS
(32) Ma, S; Phys Rev 1968, V165, P165
(33) Mathews, B; Biochem Biophys Acta 1975, V405, P442
(34) Mermin, N; Phys Rev A 1965, V137, P1441
(35) Osapay, K; J Phys Chem 1996, V100, P2698
(36) Papai, I; Int J Quant Chem:Quan Chem Symp 1990, V24, P29 HCAPLUS
(37) Perdew, J; Phys Rev B 1992, V45, P13244
(38) Perdew, J; Phys Rev L 1983, V51, P1884 HCAPLUS
(39) Rajagopal, A; Phys Rev B 1973, V7, P1912 HCAPLUS
(40) Rumelhart, D; Parallel Distributed Processing 1986
(41) Schmidt, A; Molecular Simulations 1994, V13, P347 HCAPLUS
(42) Schreckenback, G; J Phys Chem 1995, V99, P606
(43) Sejnowski, T; Complex Systems 1987, V1, P45
(44) Sham, L; Phys Rev B 1985, V32, P3883 HCAPLUS
(45) Stephens, P; Faraday Discuss 1995, V99, P103
(46) Stephens, P; J Phys Chem 1985, V89, P748 HCAPLUS
(47) Stephens, P; J Phys Chem 1987, V91, P1712 HCAPLUS
(48) Stephens, P; J Phys Chem 1994, V98, P11623 HCAPLUS
(49) Tawa, G; J Phys Chem 1996, V100, P1515 HCAPLUS
(50) Vignale, G; Phys Rev B 1988, V37, P10685
(51) Vignale, G; Phys Rev B 1988, V37, P2502
(52) Vignale, G; Phys Rev L 1987, V59, P2360 HCAPLUS
(53) Wade, R; J Am Chem Soc 1992, V114, P8284 HCAPLUS
(54) Williams, R; Neural Computation 1989, V1, P270

L86 ANSWER 11 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1998:576388** HCAPLUS
DN **129:287073**
TI Dictionary of interfaces in **proteins** (DIP). Data bank of complementary **molecular surface patches**
AU **Peissner, Robert; Goede, Andrean; Frommel, Cornelius**
CS Institute of Biochemistry, Medical Faculty of the Humboldt University (Charite), Berlin, D-10117, Germany
SO Journal of Molecular Biology (**1998**), 280(3), 535-550
CODEN: JMOBAK; ISSN: 0022-2836
PB Academic Press
DT Journal
LA English
CC 6-3 (General Biochemistry)

AB **Mol. surface areas** of **proteins** are responsible for selective **binding** of **ligands** and **protein-protein recognition**, and are considered the basis for specific interactions between different parts of a **protein**. This basic principle leads us to study the interfaces within **proteins** as a learning set for intermol. **recognition** processes of **ligands** like substrates, coenzymes, etc., and for prediction of contacts occurring during **protein folding** and assocn. For this purpose, we defined interfaces as pairs of matching **mol. surface patches** between neighboring **secondary structural** elements. All such interfaces from known **protein structures** were collected in a comprehensive data bank of interfaces in **proteins** (DIP). The up-to-date DIP contains interface files for 351 selected Brookhaven **Protein** Data Bank entries with a total of about 160,000 **surface** elements formed by 12,475 **secondary structures**. For special purposes, the inclusion of addnl. **structures** or selection of subgroups of **proteins** can be performed in an easy and straightforward manner. At. **coordinates** of the constituents of **mol. surface patches** are directly accessible as well as the corresponding contact distances from given atoms to their neighboring **secondary structural** elements. As a rule, independent of the type of **secondary structure**, the **mol. surface patches** of the **secondary structural** elements can be described as quite flat bodies with a length to width to depth ratio of about 3:2:1 for **patches** consisting of more than ten atoms. The relative orientation between two docking **patches** is strongly restricted, due to the narrow distribution of the distances between their centers of mass and of the **angles** between their normal lines, resp. The existing retrieval system for the DIP allows selection (out of the set of **mol. patches**) according to different criteria, such as geometric features, at. compn., type of **secondary structure**, contacts, etc. A fast, sequence-independent **3-D** superposition procedure was developed for automatic searches for geometrically similar **surface areas**. Using this procedure, we found a large no. of **structurally** similar interfaces of up to 30 atoms in completely unrelated **protein structures**. (c) 1998 Academic Press.

ST **protein** DIP **database folding** **conformation** docking

IT **Databases**
Protein folding
(dictionary of interfaces in **proteins** (DIP))

IT **Proteins**, general, properties
RL: PEP (Physical, engineering or chemical process); PRP (Properties); PROC (Process)
(dictionary of interfaces in **proteins** (DIP))

IT **Conformation**
(**protein**; dictionary of interfaces in **proteins** (DIP))

RE.CNT 44 THERE ARE 44 CITED REFERENCES AVAILABLE FOR THIS RECORD

RE

(1) Abagyan, R; J Biomol Struct Dynam 1992, V6, P1045
(2) Alesker, V; Protein Eng 1996, V9, P1103 HCAPLUS
(3) Alexandrov, N; J Mol Biol 1992, V225, P5 HCAPLUS
(4) Alexandrov, N; Protein Eng 1996, V9, P727 HCAPLUS
(5) Alexandrov, N; Protein Sci 1994, V3, P866 HCAPLUS
(6) Argos, P; Protein Eng 1988, V2, P101 HCAPLUS
(7) Bernstein, F; J Mol Biol 1977, V112, P535 HCAPLUS
(8) Brooks, B; J Comput Chem 1983, V4, P187 HCAPLUS
(9) Brunger, A; Science 1987, V235, P458

(10) Bures, M; Topics Stereochem 1994, V21, P467 HCAPLUS
(11) Chothia, C; Nature 1975, V254, P304 HCAPLUS
(12) Colloc'H, N; Protein Eng 1993, V6, P377 HCAPLUS
(13) Connolly, M; Biopolymers 1986, V25, P1229 HCAPLUS
(14) Connolly, M; J Am Chem Soc 1985, V107, P1118 HCAPLUS
(15) Connolly, M; Science 1983, V221, P709 HCAPLUS
(16) Doolittle, R; Sci Am 1993, V269, P50 MEDLINE
(17) Eisenhaber, F; J Comput Chem 1993, V14, P1272 HCAPLUS
(18) Eisenhaber, F; J Comput Chem 1995, V16, P273 HCAPLUS
(19) Eisenhaber, F; Protein Eng 1996, V9, P1121 HCAPLUS
(20) Fasman, K; Nucl Acids Res 1994, V22, P3462 HCAPLUS
(21) Fischer, D; Protein Eng 1995, V8, P981 HCAPLUS
(22) Gerstein, M; J Mol Biol 1995, V249, P955 HCAPLUS
(23) Goede, A; J Comput Chem 1997, V18, P1113 HCAPLUS
(24) Good, A; J Comput Aided Mol Des 1995, V9, P1 HCAPLUS
(25) Grindley, H; J Mol Biol 1993, V229, P707 HCAPLUS
(26) Hobohm, U; Protein Sci 1994, V3, P522 HCAPLUS
(27) Holm, L; Protein Sci 1992, V1, P1691 HCAPLUS
(28) Hubbard, S; Curr Opin Biotechnol 1995, V6, P375 HCAPLUS
(29) Hubbard, S; Protein Eng 1994, V7, P613 HCAPLUS
(30) Huysmans, M; Proteins:Struct Funct Genet 1991, V11, P59 HCAPLUS
(31) Janin, J; J Biol Chem 1990, V265, P16027 HCAPLUS
(32) Jones, S; Proc Natl Acad Sci USA 1996, V93, P13 HCAPLUS
(33) Jones, S; Prog Biophys Mol Biol 1995, V63, P31 HCAPLUS
(34) Kabsch, W; Biopolymers 1983, V22, P2577 HCAPLUS
(35) Kuhl, F; J Comput Chem 1984, V5, P24 HCAPLUS
(36) Lawrence, C; Proteins:Struct Funct Genet 1987, V2, P153 HCAPLUS
(37) Lessel, U; Protein Eng 1994, V7, P1175 HCAPLUS
(38) Miller, S; Protein Eng 1989, V3, P77 HCAPLUS
(39) Mizuguchi, K; Protein Eng 1995, V8, P353 HCAPLUS
(40) Orengo, C; Protein Eng 1993, V6, P485 HCAPLUS
(41) Orengo, C; Structure 1997, V5, P1093 HCAPLUS
(42) Padlan, E; Proteins:Struct Funct Genet 1990, V7, P112 HCAPLUS
(43) Pascarella, S; Protein Eng 1992, V5, P121 HCAPLUS
(44) Peng, Z; Philos Trans Roy Soc ser B 1995, V348, P43 HCAPLUS

L86 ANSWER 12 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1998:534531** HCAPLUS
DN **129:227227**
TI Flexible docking allowing induced fit in **proteins**: insights from an open to closed **conformational** isomers
AU Sandak, Bilha; Wolfson, Haim J.; Nussinov, Ruth
CS Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel
SO Proteins: Structure, Function, and Genetics (**1998**), 32(2), 159-174
CODEN: PSFGEY; ISSN: 0887-3585
PB Wiley-Liss, Inc.
DT Journal
LA English
CC 6-3 (General Biochemistry)
AB Here we dock a **ligand** onto a receptor **surface** allowing hinge-bending domain/substructural movements. Our approach mimics and manifests induced fit in **mol. recognition**. All **angular** rotations are allowed on the one hand, while a **conformational** space search is avoided on the other. Rather than dock each of the **mol.** parts sep. with subsequent reconstruction of the consistently docked **mols.**, all parts are docked simultaneously while still utilizing the position of the hinge from the start. Like pliers closing on a screw, the receptor automatically closes on its **ligand** in the best **surface**-matching way. Movements are allowed either in the **ligand** or in the larger receptor, hence reproducing induced **mol.** fit. Hinge bending

movements are frequently obsd. when **mols.** assoc. There are numerous examples of open vs. closed **conformations** taking place upon **binding.** Such movements are obsd. when the substrate **binds** to its resp. **enzyme.** In particular, such movements are of interest in allosteric **enzymes.** The movements can involve entire domains, subdomains, loops, (other) **secondary structure** elements, or between any groups of atoms connected by flexible joints. We have implemented the hinges at points and at bonds. By allowing 3-dimensional (**3-D**) rotation at the hinge, several rotations about (consecutive or nearby) bonds are implicitly taken into account. Alternatively, if required, the point rotation can be restricted to bond rotation. Here we illustrate this hinge-bending docking approach and the insight into flexibility it provides on a complex of the calmodulin with its M13 **ligand,** positioning the hinges either in the **ligand** or in the larger receptor. This automated and efficient method is adapted from **computer** vision and robotics. It enables utilizing entire **mol. surfaces** rather than focusing a priori on active sites. Hence, allows attaining the overall optimally matching **surfaces,** the extent and type of motions which are involved. Here we do not treat the **conformational** flexibility of side-chains or of very small pieces of the **mols.** Therefore, currently available methods addressing these issues and the method presented here, are complementary to each other, expanding the repertoire of **computational** docking tools foreseen to aid in studies of **recognition, conformational** flexibility and drug **design.**

ST calmodulins receptor **mol conformation** rotation **recognition**

IT **Molecular recognition**
Molecular rotation
(flexible docking allowing induced fit in **proteins:** insights from an open to closed **conformational** isomers)

IT Calmodulins
Receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(flexible docking allowing induced fit in **proteins:** insights from an open to closed **conformational** isomers)

IT **Conformation**
(**protein;** flexible docking allowing induced fit in **proteins:** insights from an open to closed **conformational** isomers)

RE.CNT 66 THERE ARE 66 CITED REFERENCES AVAILABLE FOR THIS RECORD

RE

(1) Anderson, B; Nature 1990, V344, P784 HCAPLUS
(2) Baker, E; Trends in Biochem Sci 1987, V12, P350 HCAPLUS
(3) Barford, D; Nature 1989, V340, P609 HCAPLUS
(4) Bennett, M; Prot Sci 1995, V4, P2455 HCAPLUS
(5) Bennett, W; CRC Crit Rev Biochem 1984, V15, P291 HCAPLUS
(6) Bernstein, B; Nature 1997, V385, P275 HCAPLUS
(7) Bernstein, F; J Mol Biol 1977, V112, P535 HCAPLUS
(8) Clark, K; J Comp Chem 1995, V16, P1210 HCAPLUS
(9) Connolly, M; J Appl Cryst 1983, V16, P548 HCAPLUS
(10) Connolly, M; Science 1983, V221, P709 HCAPLUS
(11) Desjarlais, R; J Med Chem 1986, V29, P2149 HCAPLUS
(12) Dobson, C; Nature 1990, V348, P198 MEDLINE
(13) Faber, H; Nature 1990, V348, P263 HCAPLUS
(14) Gerstein, M; Biochemistry 1994, V33(22), P6739 HCAPLUS
(15) Gerstein, M; J Mol Biol 1993, V234, P357 HCAPLUS
(16) Gerstein, M; J Mol Biol 1993, V229, P494 HCAPLUS
(17) Ghose, A; J Comp Chem 1985, V6, P350 HCAPLUS
(18) Goodsell, D; Proteins 1990, V8, P195 HCAPLUS
(19) Huber, R; Biochem Soc Trans 1987, V15, P1009 HCAPLUS

(20) Ikura, M; Science 1992, V256, P632 HCAPLUS
(21) Janin, J; Prog Biophys Mol Biol 1983, V42, P21 HCAPLUS
(22) Jones, G; J Mol Biol 1995, V245, P43 HCAPLUS
(23) Jones, G; J Mol Biol 1997, V267, P727 HCAPLUS
(24) Jorgensen, W; Science 1991, V254, P954 HCAPLUS
(25) Joseph, D; Science 1990, V249, P1425 HCAPLUS
(26) Judson, R; J Comp Chem 1995, V16, P1405 HCAPLUS
(27) Knegtel, R; J Mol Biol 1997, V266, P424 HCAPLUS
(28) Koshland, D; Proc Natl Acad Sci USA 1958, V44, P98 HCAPLUS
(29) Kuntz, I; J Mol Biol 1982, V161, P269 HCAPLUS
(30) Leach, A; J Comp Chem 1992, V13(6), P730 HCAPLUS
(31) Leach, A; J Mol Biol 1994, V235, P345 HCAPLUS
(32) Lin, S; J Mol Graphics 1996, V14, P78 HCAPLUS
(33) Lin, S; Proteins 1994, V18, P94 HCAPLUS
(34) Maiorov, V; Proteins 1997, V27, P410 HCAPLUS
(35) Mattevi, A; Curr Opin Struct Biol 1996, V6, P824 HCAPLUS
(36) Meador, W; Science 1992, V257, P1251 HCAPLUS
(37) Meador, W; Science 1993, V262, P1718 HCAPLUS
(38) Miller, M; Science 1989, V246, P1149 HCAPLUS
(39) Mizutani, M; J Mol Biol 1994, V243, P310 HCAPLUS
(40) Oh, B; J Bio Chem 1993, V268, P11348 HCAPLUS
(41) Paulsen, M; Proteins 1995, V21, P237 HCAPLUS
(42) Perutz, M; Nature 1970, V228, P727
(43) Rarey, M; J Mol Biol 1996, V261, P470 HCAPLUS
(44) Richards, F; Ann Rev Biophys Bioeng 1977, V6, P151 HCAPLUS
(45) Rigoutsos, D; Flexible 3D-substructure matching & novel conformer derivation in very large databases of 3D-molecular information 1996, RC 20497 90854
(46) Rini, J; Science 1992, V255, P959 HCAPLUS
(47) Sandak, B; Computer Applications in the Biosciences 1995, V11, P87 HCAPLUS
(48) Sandak, B; Journal of Biomolecular Structure & Dynamics, Proceedings of the Ninth Conversation 1996, V1, P233 HCAPLUS
(49) Sandak, B; PhD thesis, Computer Science Department, Tel Aviv University 1997
(50) Schultz, G; Curr Opin Struct Biol 1991, V1, P883
(51) Schulz, G; J Mol Biol 1990, V213, P627 HCAPLUS
(52) Sharff, A; Biochemistry 1992, V31, P10657 HCAPLUS
(53) Shoichet, B; J Comp Chem 1992, V13, P380 HCAPLUS
(54) Shoichet, B; J Mol Biol 1991, V221, P327 MEDLINE
(55) Shoichet, B; Protein Engineering 1993, V6, P723 HCAPLUS
(56) Spurlino, J; J Bio Chem 1991, V266, P5202 HCAPLUS
(57) Stanfield, R; Science 1990, V248, P712 HCAPLUS
(58) Strater, N; Proc Natl Acad Sci USA 1996, V93, P3330 MEDLINE
(59) Theriault, Y; Nature 1993, V361, P88 HCAPLUS
(60) Weber, C; Biochemistry 1991, V30, P6563 HCAPLUS
(61) Weber, I; J Mol Biol 1987, V198, P311 HCAPLUS
(62) Wedemayer, G; Science 1997, V276, P1665 HCAPLUS
(63) Welch, W; Chem Biol 1996, V3, P449 HCAPLUS
(64) Wolfson, H; Pattern Recognition Letters 1991, V12(9), P565
(65) Wriggers, W; Proteins 1997, V29, P1 HCAPLUS
(66) Yasri, A; Prot Eng 1996, V9, P959 HCAPLUS

L86 ANSWER 13 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1997:802613 HCAPLUS
DN 128:137767
TI Mode matches and their locations in the hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions
AU Mandell, Arnold J.; Selz, Karen A.; Shlesinger, Michael F.
CS Cielo Inst., Asheville, NC, 28804, USA
SO Proceedings of the National Academy of Sciences of the United States of America (**1997**), 94(25), 13576-13581
CODEN: PNASA6; ISSN: 0027-8424
PB National Academy of Sciences

DT Journal
LA English
CC 6-3 (General Biochemistry)
AB Patterns in sequences of **amino acid** hydrophobic free energies predict **secondary structures** in **proteins**. In **protein folding**, matches in hydrophobic free energy statistical wavelengths appear to contribute to selective aggregation of **secondary structures** in "hydrophobic zippers.". In a similar setting, the use of Fourier anal. to characterize the dominant statistical wavelengths of **peptide ligands'** and receptor **proteins'** hydrophobic modes to predict such matches has been limited by the aliasing and end effects of short **peptide** lengths, as well as the broad-band, mode multiplicity of many of their frequency (power) spectra. In addn., the sequence locations of the matching modes are lost in this transformation. We make new use of three techniques to address these difficulties: (i) eigenfunction construction from the linear decompn. of the lagged covariance matrixes of the **ligands** and receptors as hydrophobic free energy sequences; (ii) max. entropy, complex poles power spectra, which select the dominant modes of the hydrophobic free energy sequences or their eigenfunctions; and (iii) discrete, best bases, trigonometric wavelet transformations, which confirm the dominant spectral frequencies of the eigenfunctions and locate them as (abs. valued) moduli in the **peptide** or receptor sequence. The leading eigenfunction of the covariance matrix of a transmembrane receptor sequence locates the same transmembrane segments seen in n-block-averaged hydropathy plots while leaving the remaining hydrophobic modes unsmoothed and available for further analyses as **secondary** eigenfunctions. In these receptor eigenfunctions, we find a set of statistical wavelength matches between **peptide ligands** and their G-**protein** and tyrosine kinase coupled receptors, ranging across examples from 13.10 **amino acids** in acid fibroblast growth factor to 2.18 residues in corticotropin releasing factor. We find that the wavelet-located receptor modes in the extracellular loops are compatible with studies of receptor chimeric exchanges and point mutations. A nonbinding corticotropin-releasing factor receptor mutant is shown to have lost the signatory mode common to the normal receptor and its **ligand**. Hydrophobic free energy eigenfunctions and their transformations offer new quant. phys. homologies in **database** searches for **peptide**-receptor matches.
ST receptor **peptide ligand binding** hydrophobic zippering; **amino acid** hydrophobicity free energy **binding**
IT Free energy
(hydrophobic, of **amino acids**; mode matches and their locations in hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions)
IT **Amino acids**, biological studies
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(hydrophobic; mode matches and their locations in hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions)
IT Entropy
Hydrophobic force
Molecular association
Molecular modeling
(mode matches and their locations in hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions)
IT G **protein**-coupled receptors
Ligands
Peptides, biological studies

Receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(mode matches and their locations in hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions)

IT **Protein motifs**
(transmembrane segment; mode matches and their locations in hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions)

IT Receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(tyrosine kinase-coupled; mode matches and their locations in hydrophobic free energy sequences of **peptide ligands** and their receptor eigenfunctions)

RE.CNT 73 THERE ARE 73 CITED REFERENCES AVAILABLE FOR THIS RECORD

RE

(1) Anfinsen, C; Science 1973, V181, P223 HCAPLUS
(2) Araga, S; Proc Natl Acad Sci USA 1993, V90, P8747 HCAPLUS
(3) Beattie, J; Eur J Biochem 1996, V239, P479 HCAPLUS
(4) Bell, G; Trends Neurosci 1993, V16, P34 HCAPLUS
(5) Benya, R; Mol Pharmacol 1994, V46, P495 HCAPLUS
(6) Broomhead, D; J Phys A 1987, V20, PL563
(7) Broomhead, D; Physica D 1986, V20, P217
(8) Burg, J; Geophysics 1972, V37, P375
(9) Chen, Y; FEBS Lett 1994, V347, P279 HCAPLUS
(10) Chothia, C; Annu Rev Biochem 1984, V53, P537 HCAPLUS
(11) Daubechies, I; SIAM J Math Anal 1991, V22, P554
(12) de Gennes, P; Scaling Concepts in Polymer Physics 1985
(13) Degli, E; Eur J Biochem 1990, V190, P207
(14) Di Marzio, E; J Chem Phys 1985, V95, P1189
(15) Dill, K; Proc Natl Acad Sci USA 1993, V90, P1942 HCAPLUS
(16) Dohlman, H; Annu Rev Biochem 1991, V60, P653 HCAPLUS
(17) Fantal, W; Annu Rev Biochem 1993, V62, P453
(18) Fukuda, K; FEBS Lett 1995, V373, P177 HCAPLUS
(19) Gabor, D; J Inst Electron Eng 1946, V93, P429
(20) Gether, U; Nature (London) 1993, V362, P345 HCAPLUS
(21) Golub, G; Matrix Computations 1993, P409
(22) Gronenbron, A; Science 1994, V263, P536
(23) Hamer, R; Brain Topography 1990, V3, P43
(24) Henderson, R; J Mol Biol 1990, V213, P899 HCAPLUS
(25) Israelachvili, J; Nature (London) 1996, V379, P219 HCAPLUS
(26) Kaiser, E; Proc Natl Acad Sci USA 1983, V80, P1137 HCAPLUS
(27) Kaiser, E; Science 1984, V223, P249 HCAPLUS
(28) Kamtekar, S; FASEB J 1995, V9, P1013 HCAPLUS
(29) Kramer, R; J Biol Chem 1994, V269, P8708 HCAPLUS
(30) Kyte, J; J Mol Biol 1982, V157, P105 HCAPLUS
(31) Lesk, A; J Mol Biol 1980, V136, P225 HCAPLUS
(32) Liapakis, G; J Biol Chem 1996, V271, P20331 HCAPLUS
(33) Lin, Y; J Biol Chem 1996, V271, P5305 HCAPLUS
(34) Lumry, R; Methods Enzymol 1995, V259, P1
(35) Madan, R; Maximum Entropy and Bayesian Methods 1993, P49
(36) Manavalen, P; Nature (London) 1978, V275, P673
(37) Mandell, A; Ann N Y Acad Sci 1987, V504, P88 HCAPLUS
(38) Mandell, A; Annu Rev Pharmacol Toxicol 1984, V24, P237 HCAPLUS
(39) Mandell, A; Chaos 1997, V7, P67
(40) Mandell, A; Dynamic Patterns in Complex Systems 1988, P219
(41) Meng, F; J Biol Chem 1995, V270, P12730 HCAPLUS
(42) Metzger, T; FEBS Lett 1995, V375, P1 HCAPLUS
(43) Meunier, J; Nature (London) 1995, V377, P532 HCAPLUS
(44) Nehring, R; DNA Cell Biol 1995, V14, P939 HCAPLUS
(45) Nozaki, Y; J Biol Chem 1971, V246, P2211 HCAPLUS

(46) Ostrowski, J; Annu Rev Pharmacol Toxicol 1992, V32, P167 HCAPLUS
(47) Ozenberger, B; Mol Pharmacol 1995, V47, P82 HCAPLUS
(48) O'Dowd, B; Annu Rev Neurosci 1989, V12, P67 HCAPLUS
(49) Pashley, R; Science 1985, V229, P1088 HCAPLUS
(50) Priestly, M; Spectral Analysis and Time Series 1981, P600
(51) Ramachandran, G; Adv Protein Chem 1968, V28, P283
(52) Reich-Slotkey, R; J Biol Chem 1995, V270, P29813
(53) Reynolds, J; Proc Natl Acad Sci USA 1974, V71, P2925 HCAPLUS
(54) Rose, G; Nature (London) 1978, V272, P586 HCAPLUS
(55) Ross, P; Biochem Biophys Res Commun 1994, V205, P1836 HCAPLUS
(56) Sautel, M; Mol Pharmacol 1996, V50, P285 HCAPLUS
(57) Strader, C; Annu Rev Biochem 1994, V63, P101 HCAPLUS
(58) Strnad, J; Biochem Biophys Res Commun 1995, V216, P913 HCAPLUS
(59) Tseng, M; J Biol Chem 1995, V270, P17884 HCAPLUS
(60) Tseng, M; J Biol Chem 1995, V270, P18858 HCAPLUS
(61) Wagner, R; Nature (London) 1995, V378, P690 HCAPLUS
(62) Walker, P; J Biol Chem 1994, V269, P2863 HCAPLUS
(63) Walters, P; Symbolic Dynamics and Its Applications 1991
(64) Wang, J; J Biol Chem 1994, V269, P25966 HCAPLUS
(65) Wang, W; Proc Natl Acad Sci USA 1995, V92, P12436 HCAPLUS
(66) Watson, B; Neurobiol Dis 1996, V3, P87 HCAPLUS
(67) White, S; Biophys J 1990, V57, P911 HCAPLUS
(68) Wickerhauser, M; Adapted Wavelet Analysis from Theory to Software 1994
(69) Xue, J; J Biol Chem 1994, V269, P30195 HCAPLUS
(70) Zadina, J; Nature (London) 1997, V386, P499 HCAPLUS
(71) Zamyatnin, A; Prog Biophys Mol Biol 1972, V24, P107 MEDLINE
(72) Zetin, F; Endocrinology 1988, V122, P1121
(73) Zimmer, Y; J Biol Chem 1995, V268, P7899

L86 ANSWER 14 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1997:801238** HCAPLUS
DN **128:164152**
TI De novo heme **proteins** from **designed** combinatorial libraries
AU Rojas, Nina R. L.; Kamtekar, Satwik; Simons, Cyrena T.; Mclean, Jeremy E.; Vogel, Kathleen M.; Spiro, Thomas G.; Farid, Ramy S.; Hecht, Michael H.
CS Department of Chemistry, Princeton University, Princeton, NJ, 08544, USA
SO Protein Science (**1997**), 6(12), 2512-2524
CODEN: PRCIEI; ISSN: 0961-8368
PB Cambridge University Press
DT Journal
LA English
CC 6-3 (General Biochemistry)
AB The authors previously reported the **design** of a library of de novo **amino acid** sequences targeted to **fold** into four-helix bundles. The **design** of these sequences was based on a "binary code" strategy, in which the patterning of polar and nonpolar **amino acids** is specified explicitly, but the exact identities of the side chains is varied extensively (Kamtekar S, Schiffer JM, Xiong H, Babik JM, Hecht MH, 1993, Science 262: 1680-1685). Because of this variability, the resulting collection of **amino acid** sequences may include de novo **proteins** capable of **binding** biol. important cofactors. To probe for such **binding**, the de novo sequences were **screened** for their ability to **bind** the heme cofactor. Among an initial collection of 30 binary code sequences, 15 are shown to **bind** heme and form bright red complexes. Characterization of several of these de novo heme **proteins** demonstrated that their absorption spectra and resonance Raman spectra resemble those of natural cytochromes. Because the **design** of these sequences is based on global features of polar/nonpolar patterning, the finding that half of them **bind** heme highlights the power of the binary code strategy, and demonstrates that isolating de novo heme **proteins** does not require explicit

design of the cofactor **binding** site. Because bound heme plays a key role in the functions of many natural **proteins**, these results suggest that binary code sequences may serve as initial prototypes for the development of large collections of functionally active de novo **proteins**.

ST heme **protein design** combinatorial library

IT Absorption spectra
Combinatorial library
Dissociation constant
Molecular association
Molecular modeling
Peptide library
Protein engineering
Protein sequences
Resonance Raman spectra
Thermal stability
(characterization of heme-**binding proteins** from **designed** combinatorial libraries)

IT Cytochromes
Hemoproteins
RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process)
(characterization of heme-**binding proteins** from **designed** combinatorial libraries)

IT **Structure**-activity relationship
(**ligand-binding**, heme-**binding**; characterization of heme-**binding proteins** from **designed** combinatorial libraries)

IT **Secondary structure**
(**protein**; characterization of heme-**binding proteins** from **designed** combinatorial libraries)

IT 153224-44-3P 153224-45-4P 153224-46-5P 202759-01-1P 202759-02-2P
202759-03-3P 202759-04-4P 202759-05-5P 202759-06-6P 202759-07-7P
202759-08-8P 202759-09-9P 202759-10-2P 202759-11-3P 202759-12-4P
202759-13-5P 202759-14-6P 202759-15-7P 202759-16-8P 202759-17-9P
202759-18-0P 202759-19-1P 202759-20-4P 202759-21-5P 202759-22-6P
202759-23-7P 202759-24-8P 202759-25-9P 202759-26-0P 202759-27-1P
RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process)
(**amino acid** sequence; characterization of heme-**binding proteins** from **designed** combinatorial libraries)

IT 493-90-3, Mesoporphyrin IX 14875-96-8, Heme
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PROC (Process)
(characterization of heme-**binding proteins** from **designed** combinatorial libraries)

RE.CNT 42 THERE ARE 42 CITED REFERENCES AVAILABLE FOR THIS RECORD

RE

(1) Antonini, E; Hemoglobin and myoglobin in their reactions with ligands 1971
(2) Choi, S; J Am Chem Soc 1982, V104, P4345 HCAPLUS
(3) Choma, C; J Am Chem Soc 1994, V116, P856 HCAPLUS
(4) Desbois, A; Biochemistry 1989, V1989, P8011
(5) Dickerson, R; The enzymes:XI Oxidation-reduction, part A 1975, P397 HCAPLUS
(6) Dolphin, D; The porphyrins VII: Biochemistry, part B 1979
(7) Dumont, M; Biochemistry 1994, V33, P7368 HCAPLUS
(8) Efimov, A; Curr Opin Struct Biol 1993, V3, P379 HCAPLUS
(9) Finzel, B; J Mol Biol 1985, V186, P627 HCAPLUS
(10) Gibney, B; J Am Chem Soc 1997, V119, P2323 HCAPLUS
(11) Hamada, K; J Mol Biol 1995, V247, P947 HCAPLUS
(12) Hargrove, M; Biochemistry 1996, V35, P11293 HCAPLUS

(13) Hu, S; J Am Chem Soc 1993, V115, P12446 HCAPLUS
(14) Itagaki, E; J Biol Chem 1966, V241, P3687 HCAPLUS
(15) Johnson, B; Biotechnology 1994, V12, P1357 HCAPLUS
(16) Kamtekar, S; Science 1993, V262, P1680 HCAPLUS
(17) Kamtekar, S; [Thesis], Princeton University 1995
(18) Kirchenbaum, D; Atlas of protein spectra in the ultraviolet and visible regions 1974, V2
(19) Kitagawa, T; J Am Chem Soc 1976, V98, P5169 HCAPLUS
(20) Lemberg, R; Cytochromes 1973
(21) Lorsch, J; Biochemistry 1994, V33, P973 HCAPLUS
(22) Mathews, F; J Biol Chem 1979, V254, P1699 HCAPLUS
(23) Moore, G; Cytochromes c: Evalutionary, structural and physicochemical aspects 1990
(24) Pettigrew, G; Cytochromes c: Biological aspects 1987
(25) Rabanal, F; J Am Chem Soc 1996, V118, P473 HCAPLUS
(26) Rakshit, G; Biochemistry 1974, V13, P5317 HCAPLUS
(27) Rivera, M; Biochemistry 1992, V31, P12233 HCAPLUS
(28) Roberts, B; Proc Natl Acad Sci USA 1992, V89, P2429 HCAPLUS
(29) Robertson, D; Nature 1994, V368, P425 HCAPLUS
(30) Roy, S; Folding & Design 1997, V2, P89 HCAPLUS
(31) Roy, S; J Am Chem Soc 1997, V119, P5302 HCAPLUS
(32) Sassanfar, M; Nature 1993, V364, P550 HCAPLUS
(33) Scott, J; Science 1990, V249, P386 HCAPLUS
(34) Simons, C; [Thesis], Princeton University 1996
(35) Smulevich, G; Biochemistry 1989, V28, P9905 HCAPLUS
(36) Smulevich, G; Inorganic Chemistry 1994, V33, P4629 HCAPLUS
(37) Spiro, T; J Am Chem Soc 1974, V96, P338 HCAPLUS
(38) Spiro, T; J Am Chem Soc 1976, V98, P5482 HCAPLUS
(39) Von Bodman, B; Proc Natl Acad Sci USA 1986, V83, P9443
(40) White, W; The porphyrins, Physical chemistry, part C 1978, VV
(41) Yamanaka, T; The biochemistry of bacterial cytochromes 1992
(42) Yu, A; J Am Chem Soc 1994, V116, P4117 HCAPLUS

L86 ANSWER 15 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1997:638464 HCAPLUS
DN 127:314402
TI Constrained Corticotropin-Releasing Factor Antagonists with i-(i + 3) Glu-Lys Bridges
AU Miranda, Antonio; Lahrichi, Sabine L.; Gulyas, Jozsef; Koerber, Steven C.; Craig, Anthony G.; Corrigan, Anne; Rivier, Catherine; Vale, Wylie; Rivier, Jean
CS Clayton Foundation Laboratories for Peptide Biology, Salk Institute, La Jolla, CA, 92037, USA
SO Journal of Medicinal Chemistry (**1997**), 40(22), 3651-3658
CODEN: JMCMAR; ISSN: 0022-2623
PB American Chemical Society
DT Journal
LA English
CC 1-3 (Pharmacology)
Section cross-reference(s): 2, **34**
AB Hypothesis driven and systematic **structure**-activity relationship (SAR) investigations have resulted in the development of effective central nervous system (CNS) antagonists of corticotropin (ACTH)-releasing factor (CRF) such as .alpha.-helical CRF(9-41) and analogs of our assay std. [DPhe12,Nle21,38]hCRF(12-41). On the other hand, equally potent CRF antagonists that block the hypothalamic/pituitary/adrenal (HPA) **axis** had not been described until recently. Predictive methods, physicochem. measurements (NMR spectrometry and CD spectroscopy), and SAR studies suggest that CRF and its family members (urotensins and sauvagine) assume an .alpha.-helical **conformation** when interacting with CRF receptors. To further test this hypothesis, we have systematically scanned the hCRF(9-41) or hCRF(12-41) sequences with an i-(i + 3) bridge consisting of the Glu-Xaa-Xbb-Lys scaffold which we and others had shown

could maintain or enhance .alpha.-helical **structure**. From this series we have identified seven analogs that are either equipotent to, or 3 times more potent than, the assay std.; in addn., as presented earlier, cyclo(30-33)[DPhe12,Nle21,38,Glu30,Lys33]hCRF(12-41) (astressin) is 32 times more potent than the assay std. in blocking ACTH secretion in vitro (rat pituitary cell culture assay). In vivo, astressin is also significantly more potent than earlier antagonists at reducing hypophysial ACTH secretion in intact stressed or adrenalectomized rats. Since the corresponding linear analogs that were tested are significantly less potent, our interpretation of the increased potency of the cyclic analogs is that the introduction of the side chain to side chain bridging element (Glu30---Lys33, and to a lesser extent that of Glu14---Lys17, Glu20---Lys23, Glu23---Lys26, Glu26---Lys29, Glu28---Lys31, Glu29---Lys32, and Glu33---Lys36) induces and stabilizes in the receptor environment a putative .alpha.-helical bioactive **conformation** of the fragment that is not otherwise heavily represented. The effect of the introduction of two favored substitutions [cyclo(20-23) and cyclo(30-33)] yielded the compd. with a potency 8 times that of the assay std. but actually 12 times less than expected if the effect of the two cycles had been multiplicative. These results suggest that the pituitary CRF receptor can discriminate between slightly different identifiable **conformations**, dramatically illustrating the role that **secondary** and tertiary **structures** play in modulating biol. signaling through specific **protein-ligand** interactions.

ST corticotropin releasing factor antagonist prepn **structure**; CRF receptor antagonist signaling **ligand structure**

IT Pituitary gland
(CRF receptors of; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT Endocrine system
(adrenal-hypothalamus-pituitary, blocking of, CRF antagonists for; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT Nervous system
(central; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT Corticotropin releasing factor receptors
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
(interactions with; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT Signal transduction, biological
(modulation of; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT **.alpha.-Helix**
(prepn. and **structure** of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT **Ligands**
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PNU (Preparation, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(prepn. and **structure** of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT **Molecular modeling**
(prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT **Structure**-activity relationship
(receptor-inhibiting; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT 9015-71-8, Corticotropin-releasing factor
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(analogs; prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

IT 129133-27-3P 158068-34-9P 158068-54-3P 170809-51-5P 170809-52-6P
183615-11-4P 183615-13-6P 183615-16-9P 183615-18-1P 183615-20-5P
183615-23-8P 183615-25-0P 183615-27-2P 183615-28-3P 183868-57-7P
183906-02-7P 197443-72-4P 197443-73-5P 197443-74-6P 197443-75-7P
197443-76-8P 197443-77-9P 197443-78-0P 197443-79-1P 197443-80-4P
197443-81-5P 197443-82-6P 197443-83-7P 197443-84-8P 197443-85-9P
197443-86-0P 197527-39-2P 197527-40-5P 197527-41-6P 197527-42-7P
197527-43-8P 197527-44-9P 197527-45-0P
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PNU (Preparation, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(prepn. of constrained corticotropin-releasing factor antagonists with i-(i + 3) Glu-Lys lactam bridges)

L86 ANSWER 16 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1997:499201 HCAPLUS
DN 127:157214
TI **Computer**-assisted **design** of morphogen analogs from the atomic x-ray crystallographic **coordinates** of human osteogenic **protein** OP-1
IN Keck, Peter; Griffith, Diana L.; Carlson, William D.; Rueger, David C.; Sampath, Kuber T.
PA Creative Biomolecules, Inc., USA; Brandeis University
SO PCT Int. Appl., 176 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM C07K014-51
ICS G06F017-50
CC 2-2 (Mammalian Hormones)
FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 9726277	A2	19970724	WO 1997-US1071	19970122 <--
	W: AU, CA, JP				
	RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
	CA 2244228	AA	19970724	CA 1997-2244228	19970122 <--
	AU 9722449	A1	19970811	AU 1997-22449	19970122 <--
	AU 725295	B2	20001012		
	EP 876401	A2	19981111	EP 1997-905604	19970122 <--
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
	JP 2000501744	T2	20000215	JP 1997-526303	19970122 <--
PRAI	US 1996-589552	A	19960122 <--		
	WO 1997-US1071	W	19970122 <--		

AB Methods and compns. are provided for the **computer**-assisted **design** of morphogen analogs. Practice of the invention is enabled by the use of at least a portion of the at. **coordinates** defining the 3-dimensional **structure** of human osteogenic **protein** -1 (hOP-1) as a starting point in the **design** of the morphogen analogs. In addn., the invention provides methods for producing morphogen analogs of interest, and methods for testing whether the resulting analogs mimic or agonize human OP-1-like biol. activity. A family of morphogen analog **peptides** produced by such methods is also provided.
ST morphogen **design** osteogenic **protein** crystal **structure**
IT Bone morphogenetic **proteins**
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(7; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT **Computer application**
Crystal **structure**
Drug design
Molecular modeling
(**computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT **Proteins**, specific or class
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(morphogenetic; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT **Protein** sequences
(of morphogen analogs **designed** from human osteogenic **protein** OP-1)

IT **Conformation**
(**protein**; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT 193465-55-3P 193465-56-4P 193465-57-5P 193465-58-6P 193465-59-7P
193465-60-0P 193562-94-6P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(**computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT 193562-95-7P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(dimer stabilization; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT 193562-96-8P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(enhanced water soly.; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT 193562-51-5P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(finger 1 region of OP-1; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT 193562-49-1P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(finger 2 region of OP-1; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates** of human osteogenic **protein** OP-1)

IT 193562-50-4P
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(heel region of OP-1; **computer**-assisted **design** of morphogen analogs from the at. x-ray crystallog. **coordinates**

of human osteogenic **protein** OP-1)

L86 ANSWER 17 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1997:446520 HCAPLUS
DN 127:158732
TI Voronoi cell: new method for allocation of space among atoms: elimination of avoidable errors in calculation of atomic volume and **density**
AU **Goede, A.; Preissner, R.; Froemmel, C.**
CS Institute of Biochemistry, Humboldt University, Berlin, D-10117, Germany
SO Journal of Computational Chemistry (**1997**), **18**(**9**), **1113**-1123
CODEN: JCCHDD; ISSN: 0192-8651
PB Wiley
DT Journal
LA English
CC 9-16 (Biochemical Methods)
Section cross-reference(s): 6
AB In **computing** the vol. occupied by atoms and the d. in **proteins**, one is faced with the problem of intersecting spheres. To est. either, the space between the atoms has to be divided according to the location of the atoms relative to each other. Various methods, based on Voronoi's idea of approximating the at. space by polyhedra, have been proposed for this purpose. Comparing procedures concerned with the allocation of space among distinct atoms, we observe different partitionings of space, with deviations of more than 100% for particular atoms. Furthermore, we find that the sepg. planes of different Voronoi procedures do not meet the intersection circles of covalently linked atoms. This leads to a misallocation of space of up to 7% for atom pairs that largely differ in at. size (e.g., C-H). Several **algorithms** are neg. affected by small unallocated polyhedra ("vertex error"). These effects are cumulative for a small **protein** up to a loss of some 60 .**ANG**.3 of total vol., which would correspond to the deletion of one complete residue. To overcome these errors, instead of using dividing planes between the atoms, we use curved **surfaces**, defined as the set of those geometrical loci with equal orthogonal distance to the **surfaces** of the van der Waals spheres under consideration. The proposed dividing **surface** meets not only the intersection circle of the two van der Waals spheres but also the intersection circle of the two spheres enlarged by an arbitrary value (e.g., radius of water). This hyperbolic **surface** enveloping the Voronoi cell can be easily constructed and offers the following advantages: no misallocation of vol. for atoms of different size, no vertex error, geometrically reasonable allocation of the vol. among atoms, avoidance of discontinuities between neighboring atoms, and improved applicability to water-accessible **protein surfaces**.
ST Voronoi cell space atom calcn; atomic vol **density**
IT Volume
RL: ANT (Analyte); ANST (Analytical study)
(at.; new method for allocation of space among atoms)
IT **Algorithm**
Density
(new method for allocation of space among atoms)
IT Atoms
RL: ANT (Analyte); ANST (Analytical study)
(new method for allocation of space among atoms)
IT **Proteins**, general, properties
RL: PRP (Properties)
(new method for allocation of space among atoms)

L86 ANSWER 18 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1996:590946 HCAPLUS
DN 125:269873
TI Method and system for **protein modeling**

IN Srinivasan, Subhashini; Sudarsanam, Padmanaban
PA Immunex Corporation, USA
SO U.S., 21 pp., Cont.-in-part of U.S. 5, 453, 397.
CODEN: USXXAM
DT Patent
LA English
IC ICM G01N037-00
NCL 364496000
CC 9-16 (Biochemical Methods)
FAN.CNT 3

	PATENT NO.	KIND	DATE		APPLICATION NO.	DATE	
PI	US 5557535	A	19960917		US 1994-234812	19940428	<--
	US 5453937	A	19950926		US 1993-55050	19930428	<--
	US 5884230	A	19990316		US 1996-663809	19960614	<--
PRAI	US 1993-55050		19930428	<--			
	US 1994-234812		19940428	<--			

AB A method in a **computer** system for **modeling** a three-dimensional **structure** of a **model protein** is provided. In a preferred embodiment, the **modeling** is based upon a three-dimensional **structure** of a template **protein** and an **amino acid** sequence alignment of the **model protein** and the template **protein**. The **proteins** comprise a plurality of **amino acids** having backbone atoms and side chain atoms. For each **amino acid** in the **model protein**, when the template **protein** has an **amino acid** aligned with the **amino acid** of the **model protein**, the position of each backbone atom of the **amino acid** of the **model protein** is established based on the position of a topol. equiv. backbone atom in the aligned **amino acid** of the template **protein**. The inter-at. distance constraints for each pair of atoms with an established position is generated. Finally, the position of each atom in the **model protein** is set so that the inter-at. distances are in accordance with the constraints.

ST system **protein model**
IT **Conformation** and **Conformers**
Simulation and **Modeling**, physicochemical
(method and system for **protein modeling**)
IT **Proteins**, properties
RL: PRP (Properties)
(method and system for **protein modeling**)

L86 ANSWER 19 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1996:104717** HCAPLUS
DN **124:197693**
TI The automatic search for **ligand binding** sites in **proteins** of known three-dimensional **structure** using only geometric criteria
AU Peters, Klaus P.; Fauck, Jana; **Froemmel, Cornelius**
CS Inst. Biochem., Humboldt-Univ. Berlin, Berlin, D-10115, Germany
SO Journal of Molecular Biology (**1996**), 256(1), 201-13
CODEN: JMOBAK; ISSN: 0022-2836
PB Academic
DT Journal
LA English
CC 9-16 (Biochemical Methods)
Section cross-reference(s): 6

AB The biol. function of a **protein** typically depends on the **structure** of specific **binding** sites. These sites are located at the **surface** of the **protein mol.** and are detd. by geometrical arrangements and physico-chem. properties of

tens of non-hydrogen atoms. In this paper we describe a new **algorithm** called APROPOS, based purely on geometric criteria for identifying such **binding** sites using at. **coordinates.** For the description of the **protein** shape we use an alpha-shape **algorithm** which generates a whole family of shapes with different levels of detail. Comparing shapes of different resoln. we find cavities on the **surface** of the **protein** responsible for **ligand binding.** The **algorithm** correctly locates more than 95% of all **binding** sites for **ligands** and prosthetic groups of **mol.** mass between about 100 and 2000 Da in a representative set of **proteins.** Only in very few **proteins** does the method find **binding** sites of single ions outside the active site of **enzymes.** With one exception, we observe that interfaces between subunits show different geometric features compared to **binding** sites of **ligands.** Our results clearly support the view that **protein-protein** interactions occur between flat areas of **protein surface** whereas specific interactions of smaller **ligands** take place in pockets in the **surface.**

ST **algorithm protein ligand binding** site searching

IT **Algorithm**
(APROPOS; automatic search for **ligand binding** sites in **proteins** of known three-dimensional **structure** using only geometric criteria)

IT **Proteins,** analysis
RL: ANT (Analyte); ANST (Analytical study)
(automatic search for **ligand binding** sites in **proteins** of known three-dimensional **structure** using only geometric criteria)

IT **Conformation and Conformers**
(three dimensional; automatic search for **ligand binding** sites in **proteins** of known three-dimensional **structure** using only geometric criteria)

IT **Proteins,** specific or class
RL: ANT (Analyte); ANST (Analytical study)
(FKBP (FK 506-**binding protein**), automatic search for **ligand binding** sites in **proteins** of known three-dimensional **structure** using only geometric criteria)

IT 9001-92-7, **Proteinase** 9014-01-1, Subtilisin 9024-82-2, Inorganic pyrophosphatase
RL: ANT (Analyte); ANST (Analytical study)
(automatic search for **ligand binding** sites in **proteins** of known three-dimensional **structure** using only geometric criteria)

L86 ANSWER 20 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1994:701318 HCAPLUS
DN 121:301318
TI Synthetic, stabilized, three-dimension **polypeptides**
IN Satterthwait, Arnold C., Jr.; Arrhenius, Thomas; Chiang, Lin Chang; Cabeza, Edelmina
PA Scripps Research Institute, USA
SO PCT Int. Appl., 132 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM C07K001-00
ICS C07K007-54; C07K015-00; A61K037-02; A61K039-00; A61K039-395; C12N015-00; G01N033-53
CC 34-3 (Amino Acids, Peptides, and Proteins)
Section cross-reference(s): 1, 15

FAN.CNT 2

	PATENT NO.	KIND	DATE		APPLICATION NO.	DATE	
PI	WO 9321206	A1	19931028		WO 1993-US3032	19930331	<--
	W: AU, CA, FI, JP, NO						
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE						
	AU 9339718	A1	19931118		AU 1993-39718	19930331	<--
	US 5807979	A	19980915		US 1995-456424	19950601	<--
PRAI	US 1992-866040		19920408	<--			
	US 1993-33883		19930319	<--			
	US 1988-179160		19880408	<--			
	US 1990-607645		19901029	<--			
	US 1991-746064		19910812	<--			
	WO 1993-US3032		19930331	<--			
	US 1994-224059		19940407	<--			
OS	CASREACT 121:301318						
GI							

* STRUCTURE DIAGRAM TOO LARGE FOR DISPLAY - AVAILABLE VIA OFFLINE PRINT *

AB Synthesis of three-dimensional stabilized **peptides** which mimic the three-dimensional configuration of active site of a natural, biol. active **protein** is carried out by (1) noting the three-dimensional configuration of the active site of a known biol. active **protein**, (2) noting the **amino acid** sequence and the hydrogen bonds existing between **amino acids** and that hydrogen bonds are capable of maintaining the three-dimensional configuration of the active site, and (3) producing a synthetic three-dimensional **peptide** to mimic the **structure** of the active site. The synthetic **peptide** is synthesized so as to have the same or a similar **amino acid** sequence to the **amino acid** sequence of the active site of the biol. active **polypeptide** but with the stabilizing hydrogen bonds being replaced by a bridging divalent radical selected from the group consisting of aminomethane and aminoethane acetamidinium (N)CMe:N(H+)CH2(N), (N)CMe:N(H+)CH2CH2(N) (class I hydrogen bond mimics), and carboxybutanal hydrazone (N)N:CH(CH2)3(CO) (class II hydrogen bond mimic). Said **peptides** are represented by general cyclic **peptide** formulas (I; R1, R2 = H, C1-6 alkyl; R3 = H, C1-6 alkyl, chain of **amino acids** contg. 1-2,000 **amino acids** ; aa = **amino acid**; n = 1-2,000; R4 = any atom or **mol**. group of atoms with the required electron configuration; m = 0-6) and [II; R5 = C1-6 alkoxy, PhO, naphthyloxy, benzoxy, NH2, an **amino acid** sequence contg. 1-2,000 **amino acids**; aa = **amino acid**; n = 1-2,000; m = integer, e.g. 2; X = optionally present and if present is selected from the group consisting of CH2, NH,:CH, and :NH with double bond to CHR; R6 = optionally present and if present is selected from the group consisting of H, C1-6 alkyl, (CH2)lNH2 (wherein l = 1-6) optionally connected to an **amino acid** chain contg. 1-2,000 **amino acids**]. The hydrogen bond mimic (class I) of the cyclic **peptide** I is formed by intramol. reaction of the thioimidate group [generated by treating the corresponding thioamide R1C(S)NR2CHR3(CO) with MeI] of a **peptide** (III) with the primary NH2 group. The cyclic **peptide** II are prepd. by intramol. cyclocondensation of the hydrazide group of a **peptide** (IV) with the di-Me acetal functional group, forming the other type of the hydrogen bond mimics (class II). Thus, 5 **conformationally** restricted HIV **peptides** with the hydrogen bond mimic (class II), e.g cyclic **peptide** II [(aa)n = S-I-G-P-G-R-A-F-G, m = 2, X = bond, R6 = H, R5

= Cys-NH2] (V), which is related to the V3 loop of the HIV gp120 **protein** identified as a neutralizing epitope, were prepd. by the solid phase method. V bound to 3 HIV-**binding** murine monoclonal antibodies, at least one of which protected monkey against HIV, and reacted pos. using ELISA with sera from a patient with AIDS. HIV **peptide** II [(aa)n = S-I-S-I-G-P-G-R-A-F-Y-T-G, m = 2, X = bond, R = H, R = Cys-NH] was used to isolate human Fabs from combinatorial libraries by panning and these 5 **peptides** are potential synthetic vaccines for protection against AIDS. **Conformationally** restrained malaria **peptides** corresponding to neutralizing epitopes on various stages of Plasmodium falciparum malaria were also prepd. and are useful as a multistage vaccine. Also prepd. were epidermal growth factor analogs contg. carboxybutanal hydrazone linkage (class II) as hydrogen bond mimic.

ST three dimensional stabilized **peptide**; hydrogen bond mimic aminoethane acetamidinium; carboxybutanal hydrazone hydrogen bond mimic; malaria **peptide** neutralizing epitope; HIV gp120 **peptide** neutralizing epitope; human Fabs isolation panning HIV **peptide**; multistage vaccine malaria; vaccine HIV; **conformationally** restrained HIV malaria **peptide**; cyclic **peptide** **conformationally** restrained; active site **protein** hydrogen bond mimic; cyclization thioimidate primary amine **peptide** ; Plasmodium falciparum vaccine cyclic **peptide**; epidermal growth factor cyclic **peptide** analog

IT Antigens
RL: SPN (Synthetic preparation); PREP (Preparation)
(HIV and malaria neutralizing epitope-related three-dimensional stabilized cyclic **peptides** having hydrogen bond mimics, prepn. of, for prodn. of vaccines against HIV and malaria)

IT Vaccines
(for HIV and malaria, HIV and malaria neutralizing epitope-related three-dimensional stabilized cyclic **peptides** having hydrogen bond mimics for prodn. of)

IT **Peptides**, preparation
RL: SPN (Synthetic preparation); PREP (Preparation)
(hydrogen bond mimics (aminomethane or aminoethane acetamidinium or carboxybutanal hydrazone linkage), prepn. of, by cyclization of N -[N-(dimethoxy)butyrylpeptidyl]hydrazine derivs. or S-methylthioimidates from (thioacetyl)**peptide** (aminoethyl)amides)

IT Ring closure and formation
(of N -[N-(dimethoxy)butyrylpeptidyl]hydrazine derivs. or S-methylthioimidate deriv. from N-(thioacetyl)**peptide** -N-(aminoethyl or aminomethyl)amide , cyclic **peptides** contg. hydrazone hydrogen bond mimics)

IT Plasmodium falciparum
(vaccines for, cyclic **peptides** contg. hydrogen bond mimics as)

IT Malaria
(vaccines for, three-dimensional stabilized cyclic **peptides** having hydrogen bond mimics for)

IT **Peptides**, preparation
RL: SPN (Synthetic preparation); PREP (Preparation)
(cyclo-, with hydrogen bond mimics (aminomethane or aminoethane acetamidinium or carboxybutanal hydrazone linkage), HIV or malaria neutralizing epitope-related, prepn. of, as vaccines against AIDS and malaria)

IT Sialoglycoproteins
RL: SPN (Synthetic preparation); PREP (Preparation)
(gp120env, V3 loop of, three-dimensional stabilized cyclic **peptides** having hydrogen bond mimics related to, prepn. of, as antigens for prodn. of vaccines against HIV)

IT Virus, animal

(human immunodeficiency 1, vaccines for, three-dimensional stabilized cyclic **peptides** having hydrogen bond mimics for)

IT 26386-88-9, Diphenylphosphoryl azide
RL: RCT (Reactant); RACT (Reactant or reagent)
(acyl azide formation and Curtius rearrangement of, with alanylglycine deriv.)

IT 501-53-1, Benzyloxycarbonyl chloride
RL: RCT (Reactant); RACT (Reactant or reagent)
(acylation by, of ethylenediamine)

IT 24424-99-5, Di-tert-butyl dicarbonate 28920-43-6, 9-Fluorenylmethyloxycarbonyl chloride
RL: RCT (Reactant); RACT (Reactant or reagent)
(acylation by, of ethylenediamine deriv.)

IT 1738-69-8, Leucine benzyl ester
RL: RCT (Reactant); RACT (Reactant or reagent)
(acylation of, by (dimethoxy)pentanoic acid)

IT 107-15-3, 1,2-Ethanediamine, reactions
RL: RCT (Reactant); RACT (Reactant or reagent)
(acylation of, by benzyloxycarbonyl chloride)

IT 14794-31-1, Ethyl succinyl chloride
RL: RCT (Reactant); RACT (Reactant or reagent)
(acylation of, by ethylenediamine deriv.)

IT 159002-17-2
RL: RCT (Reactant); RACT (Reactant or reagent)
(acylation of, by fluorenylmethyloxycarbonyl chloride)

IT 75-65-0, tert-Butyl alcohol, reactions
RL: RCT (Reactant); RACT (Reactant or reagent)
(addn. of, with alanylaminomethyl isocyanate deriv.)

IT 33294-53-0
RL: RCT (Reactant); RACT (Reactant or reagent)
(condensation of, with Et N-aminoglycinate deriv.)

IT 67-64-1, Acetone, reactions
RL: RCT (Reactant); RACT (Reactant or reagent)
(condensation of, with Et hydrazinoacetate)

IT 6945-92-2, Ethyl hydrazinoacetate hydrochloride
RL: RCT (Reactant); RACT (Reactant or reagent)
(condensation of, with acetone)

IT 35661-39-3, Fmoc-Ala-OH
RL: RCT (Reactant); RACT (Reactant or reagent)
(condensation of, with resin-bound aminoalanine deriv.)

IT 1142-20-7, Z-Ala-OH
RL: RCT (Reactant); RACT (Reactant or reagent)
(esterification of, with N-hydroxysuccinimide)

IT 6066-82-6, N-Hydroxysuccinimide
RL: RCT (Reactant); RACT (Reactant or reagent)
(esterification of, with alanine deriv.)

IT 107-97-1, Sarcosine
RL: RCT (Reactant); RACT (Reactant or reagent)
(esterification of, with benzyl alc.)

IT 100-51-6, Benzyl alcohol, reactions
RL: RCT (Reactant); RACT (Reactant or reagent)
(esterification of, with sarcosine)

IT 123639-61-2, Fmoc-Glu(OBzl)-OH
RL: RCT (Reactant); RACT (Reactant or reagent)
(esterification with methanol, hydrogenolysis, and chlorination with thionyl chloride)

IT 159002-15-0
RL: RCT (Reactant); RACT (Reactant or reagent)
(hydrogenolysis and acylation of, by fluorenylmethyloxycarbonyl chloride)

IT 74-88-4, Methyl iodide, reactions
RL: RCT (Reactant); RACT (Reactant or reagent)
(methylation by, of N-(thioacetyl)**peptide**, S-thioimidate

deriv. from)
IT 56-40-6, Glycine, reactions
RL: RCT (Reactant); RACT (Reactant or reagent)
(**peptide** coupling of, with alanine deriv.)
IT 2304-94-1, N-(Benzyloxycarbonyl)-.beta.-alanine
RL: RCT (Reactant); RACT (Reactant or reagent)
(**peptide** coupling of, with glycine deriv.)
IT 1738-68-7, Benzyl glycinate
RL: RCT (Reactant); RACT (Reactant or reagent)
(**peptide** coupling of, with .beta.-alanine deriv.)
IT 126166-23-2P 126166-24-3P 126166-25-4P 157411-67-1P 158965-98-1P
158966-00-8P 158966-19-9P 158966-26-8P 158966-27-9P
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(prepn. and cyclization of)
IT 158965-97-0P
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(prepn. and cyclization of, cyclic **peptide** with hydrogen bond mimic from)
IT 637-80-9P, Ethyl hydrazinoacetate 54384-05-3P, Sarcosine benzyl ester
122652-52-2P 158966-06-4P 158966-07-5P 159002-12-7P 159002-13-8P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of)
IT 158966-32-6P 158966-33-7P 158966-34-8P 158966-35-9P 158966-36-0P
158966-37-1P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as HIV gp120 **protein**-neutralizing epitope mimic for prodn. of HIV vaccine)
IT 158966-17-7P 158966-18-8P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as antigen for prodn. of malaria vaccine)
IT 158966-01-9P 158966-02-0P 159002-14-9P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as epidermal growth factor analog contg. carboxybutanal hydrazone linkage as hydrogen bond mimic)
IT 130851-23-9DP, leucine-modified resin-bound 130851-23-9P
159002-16-1DP, leucine-modified resin-bound 159002-16-1P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as intermediate for cyclic **peptide** contg. acetamidinium linkage as hydrogen bond mimic)
IT 57260-73-8P, N-(tert-Butoxycarbonyl)ethylenediamine 72080-83-2P,
N-(Benzyloxycarbonyl)ethylenediamine 77153-05-0P, N-(Benzyloxycarbonyl)-N'-(tert-butoxycarbonyl)ethylenediamine 126166-08-3P 126166-09-4P
126166-10-7P 126166-11-8P 126166-12-9P 126166-13-0P 126198-37-6P,
N-Thioacetylsarcosine
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as intermediate for cyclic **peptide** contg. aminoethane acetamidinium linkage as hydrogen bond mimic)
IT 158966-08-6P 158966-09-7P 158966-10-0DP, resin bound 158966-10-0P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as intermediate for cyclic **peptide** contg. aminoethane amidinium linkage as hydrogen bond mimic)
IT 3235-17-4P, Z-Ala-Gly-OH 3401-36-3P, Z-Ala-OSu 126166-00-5P
126166-01-6P 126166-02-7P 126166-04-9P, Benzyl N-acetylsarcosinate
126166-05-0P, Benzyl N-(thioacetyl)sarcosinate 126166-06-1P
126166-07-2P 126198-36-5P 126198-37-6P, N-(Thioacetyl)sarcosine
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as intermediate for cyclic **peptide** contg. aminomethane acetamidinium linkage as hydrogen bond mimic)
IT 6719-33-1P, 5,5-(Dimethoxy)pentanoic acid 126166-16-3P 126166-17-4P
126166-19-6P 126166-20-9P 126166-21-0P 126166-22-1P 158965-99-2P,
Ethyl N-(isopropylidene)hydrazinoacetate hydrochloride 158966-03-1DP,

resin-bound 158966-03-1P 158966-11-1P 158966-12-2DP, Rink amide resin 158966-13-3P 158966-14-4P 158966-15-5P 158966-16-6P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as intermediate for cyclic **peptide** contg. carboxybutanal hydrazone linkage as hydrogen bond mimic)

IT 30593-19-2P, Benzyl N-(benzyloxycarbonyl)-.beta.-alanylglycinate
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as intermediate for cyclic **peptide** contg. hydrogen bond mimic)

IT 158966-28-0P 158966-29-1P 158966-30-4P 158966-31-5P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as merozoite **surface protein**-1 epitope mimic for prodn. of malaria vaccine)

IT 158966-20-2P 158966-21-3P 158966-22-4P 158966-23-5P 158966-24-6P 158966-25-7P
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn. of, as pfs25 neutralizing epitope mimic for prodn. of malaria vaccine)

IT 158966-04-2DP, leucine-modified resin-bound 158966-05-3DP, leucine-modified resin-bound
RL: SPN (Synthetic preparation); PREP (Preparation)
(prepn., methylation with Me iodide, deprotection and resin-cleavage-cyclization of)

IT 23068-91-9, Methyl 5,5-(dimethoxy)pentanoate
RL: RCT (Reactant); RACT (Reactant or reagent)
(sapon. of)

L86 ANSWER 21 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1993:120490 HCAPLUS
DN 118:120490
TI A method to identify **protein** sequences that **fold** into a known three-dimensional **structure**
IN Eisenberg, David; Bowie, James U.; Luthy, Roland
PA University of California, USA
SO PCT Int. Appl., 56 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM G01N
CC 9-16 (Biochemical Methods)
Section cross-reference(s): 6, 20

FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 9301484	A2	19930121	WO 1992-US5773	19920710 <--
	W: AU, CA, JP				
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE				
	AU 9224082	A1	19930211	AU 1992-24082	19920710 <--
	US 5436850	A	19950725	US 1994-218685	19940328 <--
PRAI	US 1991-728640		19910711 <--		
	WO 1992-US5773		19920710 <--		

AB A **computer**-assisted method is disclosed for identification of **protein** sequences that **fold** into a known three-dimensional (**3D**) **structure**. The method dets. 3 key features of each residue's environment within the **structure**: (1) the total area of the residue's side-chain that is buried by other **protein** atoms, inaccessible to solvent; (2) the fraction of the side-chain area that is covered by polar atoms (O, N) or water; and (3) the local **secondary structure**. Based on these parameters, each residue position is categorized into an environment class. In this manner, a **3D protein structure** is converted into a 1D environment string. A **3D structure** profile table is then created contg. score values that

represent the frequency of finding any of the 20 common **amino acid structures** at each position of the environment string. These frequencies are detd. from a **database** of known **protein structures** and aligned sequences. Included are **3D** compatibility search examples using a **3D structure** profile for e.g. myoglobin or cyclic AMP receptor **protein**.

ST **protein** identification **computer** three dimensional **structure; conformation protein computer**

IT Water of hydration
(**amino acid** side-chain area fraction covered by, detn. of, in **computer**-assisted identification of **protein** with known three-dimensional **structure**)

IT Solvents
(**amino acid** side-chain area inaccessible to, detn. of, in **computer**-assisted identification of **protein** with known three-dimensional **structure**)

IT **Proteins**, analysis
RL: PRP (Properties)
(**conformation** of, **computer**-assisted detn. of)

IT **Computer application**
(in identification of **protein** with known three-dimensional **structure**)

IT **Conformation** and **Conformers**
(of **proteins**, three-dimensional, **computer**-assisted detn. of)

IT Atoms
(polar, **amino acid** side-chain area fraction covered by, detn. of, in **computer**-assisted identification of **protein** with known three-dimensional **structure**)

IT Escherichia coli
(ribose-**binding protein** of, three-dimensional **structure** profile for, in compatible sequence detection)

IT **Amino acids**, properties
RL: PRP (Properties)
(**structural** properties of, in **computer**-assisted identification of **protein** with three-dimensional **structure**)

IT Actins
Myoglobins
RL: ANST (Analytical study)
(three-dimensional **structure** profile for, in compatible sequence detection)

IT Receptors
RL: ANST (Analytical study)
(cAMP, three-dimensional **structure** profile for, in compatible sequence detection)

IT Mathematics
(equations, in **computer**-assisted identification of **protein** with known three-dimensional **structure**)

IT **Proteins**, specific or class
RL: ANST (Analytical study)
(ribose-**binding**, three-dimensional **structure** profile for, of Escherichia coli, in compatible sequence detection)

IT 50-69-1, Ribose
RL: ANST (Analytical study)
(**binding protein** for, three-dimensional **structure** profile for, of Escherichia coli, in compatible sequence detection)

IT 60-92-4, Cyclic-AMP
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(receptor, three-dimensional **structure** profile for, in

compatible sequence detection)

L86 ANSWER 22 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1992:545595 HCAPLUS
DN 117:145595
TI Microscopic **modeling** of **ligand** diffusion through the **protein** leghemoglobin: **computer simulations** and experiments
AU Verkhivker, Gennady; Elber, Ron; Gibson, Quentin H.
CS Dep. Chem., Univ. Illinois, Chicago, IL, 60680, USA
SO Journal of the American Chemical Society (**1992**), 114(20), 7866-78
CODEN: JACSAT; ISSN: 0002-7863
DT Journal
LA English
CC 6-3 (General Biochemistry)
AB The diffusion of carbon monoxide through lupine legHb was investigated. The potential of mean force, the transition-state theory rate const., the friction kernel, the transmission coeff., and the diffusion const. were calcd. The **computations** are based on previous exploration of the diffusion dynamics using the mean field method (LES) and on calcns. of the reaction **coordinate**. The back of the heme pocket (close to phenylalanine 44 and phenylalanine 29) is a shallow free energy min. for the dissocd. **ligand**. The min. is directly accessible (without a barrier) from the **binding** site. The barrier for escaping from the free energy min. to the CE loop is low (approx. 3 kcal/**mol**). Once the **ligand** leaves the pocket, the diffusion is barrierless. The **ligand** escapes in two steps. In the first (activated) step the **ligand** is hopping from the heme pocket to the **protein** interior, and in the **second** step it diffuses through the **protein** matrix to the **surface** of the macromol. The transition-state theory (which is appropriate for activated processes) is used for the first part of the process. For the **second** part a diffusion **model** is constructed. The calcd. friction kernel and its power spectrum strongly depend on the reaction **coordinate**. The power spectrum is consistent with previous interpretations of the diffusion dynamics. In the first step of the process the barrier is local and the power spectrum shows only high-frequency modes. In the **second** step significant coupling to low-frequency (extended) modes is obsd., and the diffusion **coordinate** is dominated by motions of the C and the G helices of the **protein**. Exptl. results for **ligand** rebinding kinetics in lupine legHb are reported. It is shown that different diat. **ligands** have an unusually fast diffusion rate in accord with theory.
ST **ligand** diffusion legHb; **simulation ligand** diffusion legHb
IT **Ligands**
RL: PEP (Physical, engineering or chemical process); PROC (Process)
(diffusion of, in legHb interior, empirical and **mol. modeling** of)
IT Legoglobins
RL: BIOL (Biological study)
(**ligand** diffusion through, empirical and **mol. modeling** of)
IT Free energy
(of diffusion, of **ligand** in legHb, empirical and theor. studies of)
IT Potential energy and function
(of **ligand** diffusion in leg Hb interior, empirical and **mol. modeling** of)
IT **Molecular modeling**
(of **ligand** diffusion in legHb interior)

IT Kinetics of recombination
(geminate, of **ligand binding** to legHb, **mol**
. **modeling** of diffusion in relation to)
IT **Conformation and Conformers**
(**helical**, of legHb, motions of, **ligand** diffusion
response to)
IT 630-08-0, Carbon monoxide, biological studies
RL: PEP (Physical, engineering or chemical process); PROC (Process)
(diffusion of, in legHb interior, theor. and empirical studies of)
IT 7782-44-7, Oxygen, reactions 10102-43-9, Nitric oxide, reactions
RL: PRP (Properties)
(geminate recombination kinetics of, with legHb of lupine, **mol**
. **modeling** of diffusion in relation to)
IT 14875-96-8, Heme
RL: BIOL (Biological study)
(**ligand** diffusion from, in legHb, empirical and **mol**
. **modeling** of)

L86 ANSWER 23 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN 1992:58326 HCAPLUS
DN 116:58326
TI System and method for determining three-dimensional **structures** of **proteins**
IN Skolnick, Jeffrey
PA Scripps Clinic and Research Foundation, USA
SO PCT Int. Appl., 136 pp.
CODEN: PIXXD2
DT Patent
LA English
IC ICM G06F015-46
CC 20-5 (History, Education, and Documentation)
Section cross-reference(s): 6, 9
FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 9116683	A1	19911031	WO 1991-US2786	19910423 <--
	W: AU, CA, FI, JP, NO				
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, NL, SE				
	AU 9178837	A1	19911111	AU 1991-78837	19910423 <--
	JP 05501324	T2	19930311	JP 1991-509821	19910423 <--
PRAI	US 1990-513918		19900424 <--		
	WO 1991-US2786		19910423 <--		

AB The system comprises an input means such as a keyboard for specifying (entering) selected **amino acid** sequences and other data such as temp. and **fold** preferences, a RAM (random access memory) for storing such data, a ROM (read-only memory) with a stored program, a CRT (cathode ray tube) display unit and/or printer, an optional auxiliary disk storage device for storage of relevant data bases, and a microprocessor for processing the entered data, for **simulating**, under control of the stored program, the **folding** of the **protein** from its unfolded state to its **folded** (tertiary) state, and for displaying via the display unit (or printer) tertiary **conformations** of the **protein** in 3 dimensions. A block diagram of the system, flow charts showing its method of operation, and **computer** code are included, as are graphical illustrations of a **folding** pathway defined by a sequence as it **folds** into its native state.
ST **protein conformation computer modeling; folding protein computer modeling**
IT **Algorithm**
Computer program
(for **protein** three-dimensional **conformation** detn.)

IT **Computer application**
(in **protein** three-dimensional **conformation** detn.)
IT **Molecular modeling**
(of **protein** three-dimensional **conformation**, **computer** system for)
IT **Conformation** and **Conformers**
(**protein**, three-dimensional, detn. of, **computer** system for)
IT **Proteins**, biological studies
RL: PRP (Properties)
(three-dimensional **conformation** of, detn. of, **computer** system for)
IT **Simulation** and **Modeling**, biological
(Monte Carlo, in **protein** three-dimensional **conformation** detn. with **computer** system)
IT **Proteins**, specific or class
RL: PRP (Properties)
(globular, three-dimensional **conformation** of, detn. of, **computer** system for)

L86 ANSWER 24 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1989:402931** HCAPLUS
DN **111:2931**
TI A multiple sequence alignment **algorithm** for homologous **proteins** using **secondary structure** information and optionally keying alignments to functionally important sites
AU Henneke, Christina M.
CS Sch. Chem., Univ. Bath, Bath, BA2 7AY, UK
SO CABIOS, Computer Applications in the Biosciences (**1989**), 5(2), 141-50
CODEN: COABER; ISSN: 0266-7061
DT Journal
LA English
CC 6-3 (General Biochemistry)
Section cross-reference(s): 9
AB The programs described herein function as part of a suite of programs **designed** for pairwise alignment, multiple alignment, generation of randomized sequences, prodn. of alignment scores, and a sorting routine for anal. of the alignments produced. The sequence alignment programs penalize gaps (absences of residues) within regions of **protein secondary structure** and have the added option of fingerprinting **structurally** or functionally important **protein** residues. The multiple alignment program is based upon the sequence alignment method of Needleman and Wunsch and the multiple alignment extension of Barton and Sternberg. Application includes the feature of optionally weighting active site, monomer-monomer, **ligand** contact, or other important template residues to bias the alignment toward matching these residues. A sum-score for the alignments is introduced, which is independent of gap penalties. This score more adequately reflects the character of the alignments for a given scoring matrix than the gap-penalty-dependent total score described previously in the literature. In addn., individual **amino acid** similarity scores at each residue position in the alignments are printed with the alignment output to enable immediate quant. assessment of homol. at key sections of the aligned chains.
ST **protein** sequence alignment **conformation** **computer algorithm**
IT **Protein** sequences
(alignment of, of homologous **proteins**, **algorithm** using **secondary conformation** and functional sites for)
IT **Algorithm**
(for **protein** sequence alignment, using **secondary**

conformation and functional sites)
IT **Conformation and Conformers**
(**secondary**, of **proteins**, in **amino acid** sequence alignment **algorithm**)

L86 ANSWER 25 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1986:144994** HCAPLUS
DN **104:144994**
TI Comparative evaluation of the effectiveness of predicting the **secondary structure** of simple and **ligand** -containing **proteins**
AU Barkovskii, E. V.
CS Inst. Zool., Minsk, USSR
SO Vestsi Akademii Navuk BSSR, Seryya Biyalagichnykh Navuk (**1985**), (5), 109-10
CODEN: VABBA3; ISSN: 0002-3558
DT Journal
LA Russian
CC 9-10 (Biochemical Methods)
Section cross-reference(s): 6, **7**
AB The effectiveness of 3 previously published **algorithms** for predicting the **conformation** of 3 classes of **proteins** was investigated. The **protein** classes studied were (1) nucleotide-**binding proteins** (lactate dehydrogenase, alc. dehydrogenase, etc.), (2) **metalloproteins** (Hb, cytochromes, etc.), and (3) simple **proteins** (lysozyme, RNase, etc.). The accuracy of prediction depended on the **algorithm** used and the class of **proteins**. In general, the prediction was more accurate for simple **proteins** than for the other 2 classes.
ST **conformation** prediction **protein ligand algorithm**; nucleotide **binding protein conformation** prediction; **metalloprotein conformation** prediction; **enzyme conformation** prediction **algorithm**
IT **Enzymes**
Proteins
RL: PRP (Properties)
(**conformation** of, prediction of)
IT **Algorithm**
(for **conformational** prediction of **proteins**, accuracy of)
IT **Conformation and Conformers**
(of **proteins**, prediction of, **protein** classes in relation to)
IT **Enzymes**
Proteins
RL: PRP (Properties)
(metallo-, **conformation** of, prediction of)
IT **Enzymes**
Proteins
RL: PRP (Properties)
(nucleotide-**binding**, **conformation** of, prediction of)

L86 ANSWER 26 OF 26 HCAPLUS COPYRIGHT 2003 ACS
AN **1986:105613** HCAPLUS
DN **104:105613**
TI **Molecular** speleology: the exploration of crevices in **proteins** for prediction of **binding** sites, **design** of drugs and analysis of **surface recognition**
AU Lesk, Arthur M.
CS Lab. Mol. Biol., MRC, Cambridge, CB2 2QH, UK
SO Acta Crystallographica, Section A: Foundations of Crystallography (

1986), A42(2), 83-5
CODEN: ACACEQ; ISSN: 0108-7673
DT Journal
LA English
CC 9-10 (Biochemical Methods)
Section cross-reference(s): 1
AB A method is described for analyzing **mol. surface** complementarity, including the **binding** of **ligands** to **proteins** or the interaction of elements of **secondary structure** in **protein** interiors. A **computer** program can identify and **model mols.** that satisfy general criteria for good **binding** affinity. **Computational** tests are presented. This approach is likely to have useful application in the anal. of **surface recognition** in **proteins**, including the identification of **binding** sites, and in the **design** of drugs for specific targets, by (i) suggesting potential pharmacophores to the medicinal chemist for further **computational** anal. or lab. testing, (ii) suggestion of derivs. of a known **ligand** to enhance its affinity, or (iii) searching a data base of known drugs for a match to the predicted **ligand.**
ST **protein surface** complementarity analysis; **ligand binding protein** prediction; drug **design protein surface; computer** program **protein surface** complementarity
IT **Ligands**
RL: ANST (Analytical study)
(**binding** of, to **proteins**, method for **surface** complementarity anal. in study of)
IT Pharmaceuticals
(**design** of, method for **protein surface** complementarity anal. for)
IT **Computer** program
(for **protein surface** complementarity anal.)
IT **Surface structure**
(of **proteins**, complementarity of, method for study of)
IT **Proteins**
RL: ANST (Analytical study)
(**surface** complementarity of, method for study of)
IT **Computer application**
(to **protein surface** complementarity anal.)

=> fil medline
FILE 'MEDLINE' ENTERED AT 14:34:11 ON 09 JUL 2003

FILE LAST UPDATED: 8 JUL 2003 (20030708/UP). FILE COVERS 1958 TO DATE.

On April 13, 2003, MEDLINE was reloaded. See HELP RLOAD for details.

MEDLINE thesauri in the /CN, /CT, and /MN fields incorporate the MeSH 2003 vocabulary. See http://www.nlm.nih.gov/mesh/changes2003.html for a description on changes.

This file contains CAS Registry Numbers for easy and accurate substance identification.

=> d all tot

L125 ANSWER 1 OF 19 MEDLINE
AN **2002116670** MEDLINE
DN **21825961** PubMed ID: **11836225**
TI Prediction of 3D neighbours of molecular surface patches in proteins by

artificial neural networks.
AU Dietmann S; **Frommel C**
CS Medical Faculty of the Humboldt University, (Charite) Institute of Biochemistry, Monbijoustr. 2A, Berlin D-10117, Germany.. dietmann@ebi.ac.uk
SO BIOINFORMATICS, (2002 Jan) 18 (1) 167-74.
Journal code: 9808944. ISSN: 1367-4803.
CY England: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200206
ED Entered STN: 20020220
Last Updated on STN: 20020611
Entered Medline: 20020610
AB MOTIVATION: Molecular Surface Patches (MSPs) of proteins are responsible for selective interactions between internal parts of one protein molecule or between protein and other molecules. The prediction of the neighbours of a distinct Secondary Structural Element (SSE) would be an important step for protein structure prediction. RESULTS: Based on a computational analysis of complementary molecular patches of SSEs, feed-forward Neural Networks (NNs) are trained on a large set of helices for predicting the neighbours of given MSPs. Accuracy of prediction is 96% if only two types of neighbours: solvent or protein are considered, yet drops to 81% for three types of neighbours: (1) solvent, (2) helix/strand or (3) coil. Implications of the method for the prediction of protein structure and subunit interaction are discussed. As a special test case, the structurally equivalent helices of monomeric myoglobin and the homologous subunits of tetrameric haemoglobin are compared.
CT Check Tags: Comparative Study; Human; Support, Non-U.S. Gov't
Computational Biology
Hemoglobins: CH, chemistry
Models, Molecular
Myoglobin: CH, chemistry
*Neural Networks (Computer)
Protein Conformation
Protein Structure, Secondary
Protein Subunits
*Proteins: CH, chemistry
Surface Properties
CN 0 (Hemoglobins); 0 (Myoglobin); 0 (Protein Subunits); 0 (Proteins)

L125 ANSWER 2 OF 19 MEDLINE
AN **2002048829** MEDLINE
DN **21632016** PubMed ID: **11776292**
TI Matching organic libraries with protein-substructures.
AU **Preissner R**; **Goede A**; Rother K; Osterkamp F; Koert U; **Froemmel C**
CS Institute of Biochemistry, Charite, Medical Faculty of the Hunboldt-University, Berlin, Germany.
SO JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN, (2001 Sep) 15 (9) 811-7.
Journal code: 8710425. ISSN: 0920-654X.
CY Netherlands
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200206
ED Entered STN: 20020125
Last Updated on STN: 20020611
Entered Medline: 20020610
AB We present a general approach which allows automatic identification of sub-structures in proteins that resemble given three-dimensional templates. This paper documents its success with non-peptide templates

such as beta-turn mimetics. We considered well-tested turn-mimetics such as the bicyclic turned dipeptide (BTD), spiro lactam (Spiro) and the 2,5-disubstituded tetrahydrofuran (THF), a new furan-derivative which was recently developed and characterized. The detected geometric similarity between the templates and the protein patches corresponds to r.m.s.-values of 0.3 A for more than 80% of the constituting atoms, which is typical for active site comparisons of homologous proteins. This fast automatic procedure might be of biomedical value for finding special mimicking leads for particular protein sub-structures as well as for template-assembled synthetic protein (TASP) design.

CT Computer Simulation
Databases, Protein
*Drug Design
Models, Molecular
*Molecular Mimicry
Molecular Structure
Peptide Library
Protein Binding
*Proteins: CH, chemistry

CN 0 (Peptide Library); 0 (Proteins)

L125 ANSWER 3 OF 19 MEDLINE

AN **2000304908** MEDLINE

DN **20304908** PubMed ID: **10843865**

TI Conservation of substructures in proteins: interfaces of secondary structural elements in proteasomal subunits.

AU Gille C; **Goede A; Preissner R;** Rother K; **Frommel C**

CS Institute of Biochemistry Charite, Medical Faculty of the Humboldt University, Monbijoustrasse 2a, Berlin, 10117, Germany.. christoph.gille@charite.de

SO JOURNAL OF MOLECULAR BIOLOGY, (2000 Jun 16) 299 (4) 1147-54. Journal code: 2985088R. ISSN: 0022-2836.

CY ENGLAND: United Kingdom

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS Priority Journals; Space Life Sciences

EM 200007

ED Entered STN: 20000720
Last Updated on STN: 20000720
Entered Medline: 20000712

AB It is observed that during divergent evolution of two proteins with a common phylogenetic origin, the structural similarity of their backbones is often preserved even when the sequence similarity between them decreases to a virtually undetectable level. Here we analyzed, whether the conservation of structure along evolution involves also the local atomic structures in the interfaces between secondary structural elements. We have used as study case one protein family, the proteasomal subunits, for which 17 crystal structures are known. These include 14 different subunits of Saccharomyces cerevisiae, 2 subunits of Thermoplasma acidophilum and one subunit of Escherichia coli. The structural core of the 17 proteasomal subunits has 23 secondary structural elements. Any two adjacent secondary structural elements form a molecular interface consisting of two molecular patches. We found 61 interfaces that occurred in all 17 subunits. The 3D shape of equivalent molecular patches from different proteasomal subunits were compared by superposition. Our results demonstrate that pairs of equivalent molecular patches show an RMSD which is lower than that of randomly chosen patches from unrelated proteins. This is true even when patch comparisons with identical residues were excluded from the analysis. Furthermore it is known that the sequential dissimilarity is correlated to the RMSD between the backbones of the members of protein families. The question arises whether this is also true for local atomic structures. The results show that the

correlation of individual patch RMSD values and local sequence dissimilarities is low and has a wide range from 0 to 0.41, however, it is surprising that there is a good correlation between the average RMSD of all corresponding patches and the global sequence dissimilarity. This average patch RMSD correlates slightly stronger than the C(alpha)-trace RMSD to the global sequence dissimilarity.

CT Algorithms
Amino Acid Sequence
Binding Sites
*Conserved Sequence
*Cysteine Endopeptidases: CH, chemistry
*Cysteine Endopeptidases: ME, metabolism
Databases, Factual
Escherichia coli: EN, enzymology
Evolution, Molecular
Models, Molecular
*Multienzyme Complexes: CH, chemistry
*Multienzyme Complexes: ME, metabolism
Protein Structure, Quaternary
Protein Structure, Secondary
Saccharomyces cerevisiae: EN, enzymology
Sequence Alignment
Software
Thermoplasma: EN, enzymology
CN 0 (Multienzyme Complexes); EC 3.4.22 (Cysteine Endopeptidases); EC 3.4.99.46 (multicatalytic endopeptidase complex)

L125 ANSWER 4 OF 19 MEDLINE
AN **2000169613** MEDLINE
DN **20169613** PubMed ID: **10705435**
TI Homonyms and synonyms in the Dictionary of Interfaces in Proteins (DIP).
AU **Preissner R; Goede A; Frommel C**
CS Medical Faculty, Humboldt University, Charite, Institute of Biochemistry, Berlin, Germany.. preissner@rz.hu-berlin.de
SO BIOINFORMATICS, (1999 Oct) 15 (10) 832-6.
Journal code: 9808944. ISSN: 1367-4803.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200004
ED Entered STN: 20000421
Last Updated on STN: 20000421
Entered Medline: 20000411
AB MOTIVATION: Should reports on molecular mimicry in particular cases, e.g. responsible for cross-reactivity, be considered as accidental or as a general principle in protein evolution? To answer this question, two types of similarity have to be considered: those in homologues (synonyms) and resemblance between patches from unrelated proteins (homonyms). RESULTS: All interfaces from known protein structures were collected in a comprehensive data bank [Dictionary of Interfaces in Proteins (DIP)]. A fast, sequence-independent, three-dimensional superposition procedure was developed to search automatically for geometrically similar surface areas. Surprisingly, we found a large number of structurally similar interfaces on the surface of unrelated proteins. Even patches from different types of secondary structure were found resembling each other. The putative functional meaning of homonyms is demonstrated with striking examples.
CT Check Tags: Comparative Study
Algorithms
Amino Acid Sequence
Computational Biology
*Databases, Factual

*Dictionaries, Chemical
Models, Molecular
Molecular Mimicry
Molecular Sequence Data
Protein Conformation
*Proteins: CH, chemistry
*Proteins: GE, genetics
CN 0 (Proteins)

L125 ANSWER 5 OF 19 MEDLINE
AN **2000027473** MEDLINE
DN **20027473** PubMed ID: **10556242**
TI Spare parts for helix-helix interaction.
AU **Preissner R; Goede A; Frommel C**
CS Institute of Biochemistry, Charite, Medical Faculty of the Humboldt University, Monbijoustr 2, 10117 Berlin, Germany.
SO PROTEIN ENGINEERING, **(1999 Oct)** 12 (10) 825-32.
Journal code: 8801484. ISSN: 0269-2139.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200001
ED Entered STN: 20000124
Last Updated on STN: 20000124
Entered Medline: 20000112
AB About 6000 contact regions (patches) of helix-to-helix packing from 300 well-resolved non-homologous protein structures were considered. The patches were defined by the spatial helical neighbors and were estimated in atomic detail using a variable distance criterion. The following questions are addressed. (1) Are the amino acid preferences and atomic composition of distinct types of helical patches indicative for the type of their neighbor? Distributions of size, atomic composition and packing density are compared for different types of helical interfaces. Thereby contact preferences are derived for parts of secondary structures adjoining each other or pointing towards the solvent. (2) Is it possible to cluster helical patches according to their structural similarity? For these purposes the patches were classified with an automatic sequence-independent superposition procedure which yields a distinctively reduced set of representative interfaces. On this basis, the methodology for finding exchangeable patches in different proteins is demonstrated.
CT **Databases, Factual**
Models, Molecular
Protein Folding
***Protein Structure, Secondary**
*Proteins: CH, chemistry
*Proteins: ME, metabolism
CN 0 (Proteins)

L125 ANSWER 6 OF 19 MEDLINE
AN **1999342043** MEDLINE
DN **99342043** PubMed ID: **10411900**
TI A systematic study of low-resolution recognition in protein--protein complexes.
AU Vakser I A; Matar O G; Lam C F
CS Department of Cell and Molecular Pharmacology, Medical University of South Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA.. vakseri@musc.edu
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, **(1999 Jul 20)** 96 (15) 8477-82.
Journal code: 7505876. ISSN: 0027-8424.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English

FS Priority Journals
EM 199908
ED Entered STN: 19990910
Last Updated on STN: 19990910
Entered Medline: 19990823
AB A comprehensive nonredundant database of 475 cocrystallized protein-protein complexes was used to study low-resolution recognition, which was reported in earlier docking experiments with a small number of proteins. The docking program GRAMM was used to delete the atom-size structural details and systematically dock the resulting molecular images. The results reveal the existence of the low-resolution recognition in 52% of all complexes in the database and in 76% of the 113 complexes with an interface area >4,000 A(2). Limitations of the docking and analysis tools used in this study suggest that the actual number of complexes with the low-resolution recognition is higher. However, the results already prove the existence of the low-resolution recognition on a broad scale.
CT Check Tags: Support, U.S. Gov't, Non-P.H.S.
Algorithms
Binding Sites
Crystallization
Databases
Fourier Analysis
Ligands
Models, Molecular
Protein Binding
Protein Structure, Secondary
Protein Structure, Tertiary
***Proteins: CH, chemistry**
Receptors, Cell Surface: CH, chemistry
CN 0 (**Ligands**); 0 (Proteins); 0 (Receptors, Cell Surface)

L125 ANSWER 7 OF 19 MEDLINE
AN 1999180410 MEDLINE
DN 99180410 PubMed ID: 10082373
TI Enzyme-mononucleotide interactions: three different folds share common structural elements for ATP recognition.
AU Denessiouk K A; Lehtonen J V; Johnson M S
CS Department of Biochemistry and Pharmacy, Abo Akademi University, Turku, Finland.
SO PROTEIN SCIENCE, **(1998 Aug)** 7 (8) 1768-71.
Journal code: 9211750. ISSN: 0961-8368.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199905
ED Entered STN: 19990601
Last Updated on STN: 19990601
Entered Medline: 19990519
AB Three ATP-dependent enzymes with different folds, cAMP-dependent protein kinase, D-Ala:D-Ala ligase and the alpha-subunit of the alpha2beta2 ribonucleotide reductase, have a similar organization of their ATP-binding sites. The most meaningful similarity was found over 23 structurally equivalent residues in each protein and includes three strands each from their beta-sheets, in addition to a connecting loop. The equivalent secondary structure elements in each of these enzymes donate four amino acids forming key hydrogen bonds responsible for the common orientation of the "AMP" moieties of their ATP-**ligands**. One lysine residue conserved throughout the three families binds the alpha-phosphate in each protein. The common fragments of structure also position some, but not all, of the equivalent residues involved in hydrophobic contacts with the adenine ring. These examples of convergent evolution reinforce the view that different proteins can fold in different ways to produce similar

structures locally, and nature can take advantage of these features when structure and function demand it, as shown here for the common mode of ATP-binding by three unrelated proteins.

CT Check Tags: Support, Non-U.S. Gov't
Adenine: CH, chemistry
*Adenosine Triphosphate: CH, chemistry
Allosteric Site
Computer Simulation
Cyclic AMP-Dependent Protein Kinases: CH, chemistry
Databases, Factual
***Enzymes: CH, chemistry**
Models, Molecular
*Nucleotides: CH, chemistry
Peptide Synthases: CH, chemistry
Phosphates: CH, chemistry
Protein Binding
Protein Structure, Secondary
Ribonucleotide Reductases: CH, chemistry
Ribose: CH, chemistry

RN 50-69-1 (Ribose); 56-65-5 (Adenosine Triphosphate); 73-24-5 (Adenine)
CN 0 (Enzymes); 0 (Nucleotides); 0 (Phosphates); EC 1.17.4 (Ribonucleotide Reductases); EC 2.7.1.37 (Cyclic AMP-Dependent Protein Kinases); EC 6.3.2. (Peptide Synthases); EC 6.3.2.4 (D-alanylalanine synthetase)

L125 ANSWER 8 OF 19 MEDLINE
AN **1999077019** MEDLINE
DN **99077019** PubMed ID: **9862203**
TI Sequences annotated by structure: a tool to facilitate the use of structural information in sequence analysis.
AU Milburn D; Laskowski R A; Thornton J M
CS Department of Biochemistry and Molecular Biology, University College, London, UK.
SO PROTEIN ENGINEERING, **(1998 Oct)** 11 (10) 855-9.
Journal code: 8801484. ISSN: 0269-2139.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199903
ED Entered STN: 19990316
Last Updated on STN: 20000303
Entered Medline: 19990303
AB With the aim of bridging the gap between protein sequence and structural analyses, we have developed a tool to aid the identification of new protein sequences by recognizing distant homologues using structural information. The tool generates sequence annotated by structure (SAS) files, applying structural information derived from structural analyses to a given protein sequence. A World Wide Web interface allows a given sequence to be submitted either for structural annotation or, where its structure is unknown, for search and alignment against sequences of known structure. In both cases, SAS will colour residues in the sequence of known structure according to a selection of properties, including secondary structure, interatomic contacts and active site information. SAS can also be used to inspect properties of a single structure.
CT Check Tags: Support, Non-U.S. Gov't
Amino Acid Sequence
Binding Sites
DNA-Binding Proteins: CH, chemistry
Databases, Factual
Hydrogen Bonding
***Internet**
Lactalbumin: CH, chemistry
Lactalbumin: ME, metabolism

Ligands
Metals: ME, metabolism
Molecular Sequence Data
Muramidase: CH, chemistry
Muramidase: ME, metabolism
***Protein Conformation**
Protein Folding
Protein Structure, Secondary
***Proteins: CH, chemistry**
Sequence Alignment
*Sequence Analysis
Sequence Homology, Amino Acid
Software
Structure-Activity Relationship
RN 9013-90-5 (Lactalbumin)
CN 0 (DNA-Binding Proteins); 0 (**Ligands**); 0 (Metals); 0 (Proteins); EC 3.2.1.17 (Muramidase)

L125 ANSWER 9 OF 19 MEDLINE
AN 1999000702 MEDLINE
DN 99000702 PubMed ID: 9784114
TI CoMFA-based prediction of agonist affinities at recombinant D1 vs D2 dopamine receptors.
AU Wilcox R E; Tseng T; Brusniak M Y; Ginsburg B; Pearlman R S; Teeter M; DuRand C; Starr S; Neve K A
CS Molecular Pharmacology Laboratory, College of Pharmacy, University of Texas at Austin, Austin, Texas 78712-1074, USA..
wilcoxrich@mail.utexas.edu
NC RR08579A (NCRR)
SO JOURNAL OF MEDICINAL CHEMISTRY, **(1998 Oct 22)** 41 (22) 4385-99.
Journal code: 9716531. ISSN: 0022-2623.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199811
ED Entered STN: 19990106
Last Updated on STN: 19990106
Entered Medline: 19981123
AB We have previously shown that using agonist affinity at recombinant receptors selectively expressed in clonal cells as the dependent variable in three-dimensional quantitative structure-activity relationship studies (3D-QSAR) presents a unique opportunity for accuracy and precision in measurement. Thus, a comparison of affinity's structural determinants for a set of compounds at two different recombinant dopamine receptors represents an attainable goal for 3D-QSAR. A molecular database of bound conformations of 16 structurally diverse agonists was established by alignment with a high-affinity template compound for the D1 receptor, 3-allyl-6-bromo-7,8-dihydroxy-1-phenyl-2,3,4, 5-tetrahydro-1H-benzazepin. A second molecular database of the bound conformations of the same compounds was established against a second template for the D2 receptor, bromocriptine. These aligned structures suggested three-point pharmacophore maps (one cationic nitrogen and two electronegative centers) for the two dopamine receptors, which differed primarily in the height of the nitrogen above the plane of the catechol ring and in the nature of the hydrogen-bonding region. The ln(1/KL) values for the low-affinity agonist binding conformation at recombinant D1 and D2 dopamine receptors stably expressed in C6 glioma cells were used as the target property for the CoMFA (comparative molecular field analysis) of the 16 aligned structures. The resulting CoMFA models yielded cross-validated R2 (q2) values (standard error of prediction) of 0. 879 (1.471, with five principal components) and 0.834 (1.652, with five principal components) for D1 and D2 affinity, respectively. The simple R2 values (standard error of the

estimate) were 0.994 (0.323) and 0.999 (0.116), respectively, for D1 and D2 receptor. F values were 341 and 2465 for D1 and D2 models, respectively, with 5 and 10 df. The predictive utility of the CoMFA model was evaluated at both receptors using the dopamine agonists, apomorphine and 7-OH-DPAT. Predictions of KL were accurate at both receptors. Flexible 3D searches of several chemical databases (NCI, MDDR, CMC, ACD, and Maybridge) were done using basic pharmacophore models at each receptor to determine the similarity of hit lists between the two models. The D1 and D2 models yielded different lists of lead compounds. Several of the lead compounds closely resembled high-affinity training set compounds. Finally, homology modeling of agonist binding to the D2 receptor revealed some consistencies and inconsistencies with the CoMFA-derived D2 model and provided a possible rationale for features of the D2 CoMFA contour map. Together these results suggest that CoMFA-homology based models may provide useful insights concerning differential agonist-receptor interactions at related receptors. The results also suggest that comparisons of CoMFA models for two structurally related receptors may be a fruitful approach for differential QSAR.

CT Check Tags: Animal; Human; Support, Non-U.S. Gov't; Support, U.S. Gov't, Non-P.H.S.; Support, U.S. Gov't, P.H.S.
Binding Sites
Databases, Factual
*Dopamine Agonists: CH, chemistry
Dopamine Agonists: ME, metabolism
Dopamine Agonists: PD, pharmacology
Ligands
Macaca mulatta
***Models, Molecular**
Molecular Conformation
Protein Structure, Secondary
Rats
***Receptors, Dopamine D1: AG, agonists**
Receptors, Dopamine D1: BI, biosynthesis
Receptors, Dopamine D1: CH, chemistry
Receptors, Dopamine D1: ME, metabolism
***Receptors, Dopamine D2: AG, agonists**
Receptors, Dopamine D2: BI, biosynthesis
Receptors, Dopamine D2: CH, chemistry
Receptors, Dopamine D2: ME, metabolism
Recombinant Proteins: AG, agonists
Recombinant Proteins: CH, chemistry
Recombinant Proteins: ME, metabolism
Structure-Activity Relationship
Tumor Cells, Cultured

CN 0 (Dopamine Agonists); 0 (**Ligands**); 0 (Receptors, Dopamine D1); 0 (Receptors, Dopamine D2); 0 (Recombinant Proteins)

L125 ANSWER 10 OF 19 MEDLINE
AN **1998417653** MEDLINE
DN **98417653** PubMed ID: **9743635**
TI Supersites within superfolds. Binding site similarity in the absence of homology.
AU Russell R B; Sasieni P D; Sternberg M J
CS Biomolecular Modelling Laboratory, Lincoln's Inn Fields, PO Box 123, London WC2A 3PX, UK.
SO JOURNAL OF MOLECULAR BIOLOGY, **(1998 Oct 2)** 282 (4) 903-18.
Journal code: 2985088R. ISSN: 0022-2836.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals; Space Life Sciences
EM 199907
ED Entered STN: 19990727

Last Updated on STN: 19990727
Entered Medline: 19990712

AB A method is presented to assess the significance of binding site similarities within superimposed protein three-dimensional (3D) structures and applied to all similar structures in the Protein Data Bank. For similarities between 3D structures lacking significant sequence similarity, the important distinction was made between remote homology (an ancient common ancestor) and analogy (likely convergence to a folding motif) according to the structural classification of proteins (SCOP) database. Supersites were defined as structural locations on groups of analogous proteins (i.e. superfolds) showing a statistically significant tendency to bind substrates despite little evidence of a common ancestor for the proteins considered. We identify three potentially new superfolds containing supersites: ferredoxin-like folds, four-helical bundles and double-stranded beta helices. In addition, the method quantifies binding site similarities within homologous proteins and previously identified supersites such as that found in the beta/alpha (TIM) barrels. For the nine superfolds, the accuracy of predictions of binding site locations is assessed. Implications for protein evolution, and the prediction of protein function either through fold recognition or tertiary structure comparison, are discussed.

CT Check Tags: Animal; Human; Support, Non-U.S. Gov't
***Binding Sites**
***Databases, Factual**
Evolution, Molecular
Ferredoxins: ME, metabolism
Ligands
Models, Molecular
Protein Conformation
*Protein Folding
Protein Structure, Secondary
***Proteins: CH, chemistry**
Proteins: ME, metabolism
Sequence Alignment
Sequence Homology, Amino Acid
Statistical Distributions
Structure-Activity Relationship

CN 0 (Ferredoxins); 0 (**Ligands**); 0 (Proteins)

L125 ANSWER 11 OF 19 MEDLINE
AN **1998332753** MEDLINE
DN **98332753** PubMed ID: **9665855**
TI Dictionary of interfaces in proteins (DIP). Data bank of complementary molecular surface patches.
AU **Preissner R; Goede A; Frommel C**
CS Institute of Biochemistry, Medical Faculty of the Humboldt University (Charite), Monbijoustr. 2A, Berlin, D-10117, Germany.
SO JOURNAL OF MOLECULAR BIOLOGY, **(1998 Jul 17) 280 (3) 535-50.**
Journal code: 2985088R. ISSN: 0022-2836.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199808
ED Entered STN: 19980820
Last Updated on STN: 19980820
Entered Medline: 19980807

AB Molecular surface areas of proteins are responsible for selective binding of **ligands** and protein-protein recognition, and are considered the basis for specific interactions between different parts of a protein. This basic principle leads us to study the interfaces within proteins as a

learning set for intermolecular recognition processes of **ligands** like substrates, coenzymes, etc., and for prediction of contacts occurring during protein folding and association. For this purpose, we defined interfaces as pairs of matching molecular surface patches between neighboring secondary structural elements. All such interfaces from known protein structures were collected in a comprehensive data bank of interfaces in proteins (DIP).The up-to-date DIP contains interface files for 351 selected Brookhaven Protein Data Bank entries with a total of about 160,000 surface elements formed by 12,475 secondary structures. For special purposes, the inclusion of additional structures or selection of subgroups of proteins can be performed in an easy and straightforward manner. Atomic coordinates of the constituents of molecular surface patches are directly accessible as well as the corresponding contact distances from given atoms to their neighboring secondary structural elements.As a rule, independent of the type of secondary structure, the molecular surface patches of the secondary structural elements can be described as quite flat bodies with a length to width to depth ratio of about 3:2:1 for patches consisting of more than ten atoms. The relative orientation between two docking patches is strongly restricted, due to the narrow distribution of the distances between their centers of mass and of the angles between their normal lines, respectively.The existing retrieval system for the DIP allows selection (out of the set of molecular patches) according to different criteria, such as geometric features, atomic composition, type of secondary structure, contacts, etc. A fast, sequence-independent 3-D superposition procedure was developed for automatic searches for geometrically similar surface areas. Using this procedure, we found a large number of structurally similar interfaces of up to 30 atoms in completely unrelated protein structures.

CT ***Databases, Factual**
Forecasting
Models, Molecular
Protein Folding
Protein Structure, Secondary
***Proteins: CH, chemistry**
Solvents: CH, chemistry
Surface Properties

CN 0 (Proteins); 0 (Solvents)

L125 ANSWER 12 OF 19 MEDLINE

AN **1998075920** MEDLINE

DN **98075920** PubMed ID: **9415438**

TI Applying experimental data to protein fold prediction with the genetic algorithm.

AU Dandekar T; Argos P

CS European Molecular Biology Laboratory, Heidelberg, Germany.

SO PROTEIN ENGINEERING, **(1997 Aug)** 10 (8) 877-93.
Journal code: 8801484. ISSN: 0269-2139.

CY ENGLAND: United Kingdom

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS Priority Journals

EM 199802

ED Entered STN: 19980224
Last Updated on STN: 19980224
Entered Medline: 19980211

AB Specific residue interactions as revealed from a few and readily available experiments can be quite important in shaping a protein's tertiary topology by complementing basic and general folding principles. This experimental information is employed in structure prediction (mainchain topology) based on sequence knowledge and the genetic algorithm with its ability to optimize simultaneously many parameters. Examples investigated include the distribution of cysteinyl S-S bonds, protein side-chain

ligands to iron-sulfur cages, cofactor-**ligands**, crosslinks amongst side-chains, and conserved hydrophobic and catalytic residues. Such interactions yield an improvement in the predicted topology (0.4-6.6 A root mean square deviation in the positions of the backbone C alpha-atoms relative to those observed) compared with those resulting from simulations relying only on basic protein folding principles. For several examples the resultant topology depended critically on knowledge of the few and specific interactions such that the relationship between predicted and observed C alpha-positions was near random without their use. The combined methodology (experimental data and the genetic algorithm) should prove helpful in settings where experiment and theory can cooperate in successive steps to elucidate an unknown structure.

CT ***Algorithms**
Computer Simulation
Databases, Factual
Disulfides: CH, chemistry
Evolution, Molecular
Models, Molecular
*Protein Folding
Protein Structure, Secondary
Protein Structure, Tertiary
***Proteins: CH, chemistry**
Proteins: GE, genetics

CN 0 (Disulfides); 0 (Proteins)

L125 ANSWER 13 OF 19 MEDLINE
AN **97459773** MEDLINE
DN **97459773** PubMed ID: **9315733**
TI Inverse sequence similarity in proteins and its relation to the three-dimensional fold.
AU **Preissner R; Goede A;** Michalski E; **Frommel C**
CS Institute of Biochemistry, Charite, Berlin, Germany.
SO FEBS LETTERS, (1997 Sep 8) 414 (2) 425-9.
Journal code: 0155157. ISSN: 0014-5793.
CY Netherlands
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199710
ED Entered STN: 19971105
Last Updated on STN: 19971105
Entered Medline: 19971021
AB Nowadays the most successful strategy for the prediction of the tertiary structure of proteins is the homology-based modelling using known structures. A real chance to predict the general fold of a protein arises only in cases with a sufficient sequence homology (e.g. 27% over 100 residues). In this analysis we examine the phenomenon of inverse sequence similarity (ISS) in proteins and its structural meaning. In sequence data bases we found a lot of examples for ISS up to 34% identity over 204 residues and a surprisingly large number of self-inverse protein sequences. By inspection of inverse similar sequence pairs with known tertiary structures we observe that inverse sequence alignments above the threshold indicating structural similarity generally do not imply comparable folds for both. From our analysis we conclude that the straightforward employment of ISS for protein structure prediction fails even above the known threshold for 'safe similarity'.
CT *Amino Acid Sequence
*Computer Simulation
Enzymes: CH, chemistry
Information Systems
*Models, Molecular
Molecular Sequence Data

*Protein Folding
*Protein Structure, Tertiary
*Proteins: CH, chemistry
Ribonucleases: CH, chemistry
Sequence Alignment
CN 0 (Enzymes); 0 (Proteins); EC 3.1.- (Ribonucleases)

L125 ANSWER 14 OF 19 MEDLINE
AN **96407680** MEDLINE
DN **96407680** PubMed ID: **8811733**
TI Prediction of secondary structural content of proteins from their amino acid composition alone. II. The paradox with secondary structural class.
AU Eisenhaber F; **Frommel C**; Argos P
CS Institut fur Biochemie der Charite, Medizinische Fakultat, Humboldt-Universitat zu Berlin, Berlin-Mitte, Germany.
SO PROTEINS, **(1996 Jun)** 25 (2) 169-79.
Journal code: 8700181. ISSN: 0887-3585.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199612
ED Entered STN: 19970128
Last Updated on STN: 19970128
Entered Medline: 19961205
AB The success rates reported for secondary structural class prediction with different methods are contradictory. On one side, the problem of recognizing the secondary structural class of a protein knowing only its amino acid composition appears completely solved by simply applying jury decision with an elliptically scaled distance function. Chou and coworkers repeatedly (see Crit. Rev. Biochem. Mol. Biol. 30:275-349, 1995) published prediction accuracies near 100%. On the other hand, traditional secondary structure prediction techniques achieve success rates of about 70% for the secondary structural state per residue and about 75% for structural class only with extensive input information (full sequence of the query protein, its amino acid composition and length, multiple alignments with homologous sequences). In this article, we resolve the paradox and consider (1) the question of the secondary structural class definition, (2) the role of the representativity of the test set of protein tertiary structure for the current state of the Protein Data Bank (PDB); and (3) we estimate the real impact of amino acid composition on secondary structural class. We formulate three objective criteria for a reasonable definition of secondary structural classes and show that only the criterion of Nakashima et al. (J. Biochem. 99:153-162, 1986) complies with all of them. Only this definition matches the distribution of secondary structural content in representative PDB subsets, whereas other criteria leave many proteins (up to 65% of all PDB entries) simply unassigned. We review critically specialized secondary-structural class prediction methods, especially those of Chou and coworkers, which claim almost 100% accuracy using only amino acid composition, and resolve the paradox that these prediction accuracies are better than those from secondary structure predictions from multiple alignments. We show (i) that these techniques rely on a preselection of test sets which removes irregular proteins and other proteins without any class assignment (about 35% of all PDB entries); and (ii) that even for preselected representative test sets, the success rate drops to 60% and lower for a 4-type classification (alpha, beta, alpha + beta, alpha/beta). The prediction accuracies fall to about 50% if the secondary structural class definition of Nakashima et al. is applied and only few irregular proteins are preselected and removed from automatically generated, representative subsets of the PDB. We have applied two new vector decomposition methods for secondary structural content prediction from amino acid composition alone, with and without consideration of amino acid

compositional coupling in the learning set of tertiary structures respectively, to the problem of class prediction and achieve about 60% correct assignment among four classes (alpha, beta, mixed, irregular) as well as single sequence-based secondary structure prediction methods like GORIII and COMBI. Our results demonstrate that 60% correctness is the upper limit for a 4-type class prediction from amino acid composition alone for an unknown query protein and that consideration of compositional coupling does not improve the prediction success. The prediction program SSCP offering secondary structural class assignment for query compositions and sequences has been made available as a World Wide Web and E-mail service.

CT Check Tags: Comparative Study; Support, Non-U.S. Gov't
*Amino Acids: CH, chemistry
Computer Simulation
Forecasting: MT, methods
Models, Chemical
***Protein Structure, Secondary**
CN 0 (Amino Acids)

L125 ANSWER 15 OF 19 MEDLINE
AN **96407679** MEDLINE
DN **96407679** PubMed ID: **8811732**
TI Prediction of secondary structural content of proteins from their amino acid composition alone. I. New analytic vector decomposition methods.
AU Eisenhaber F; Imperiale F; Argos P; **Frommel C**
CS Institut fur Biochemie der Charite, Medizinische Fakultat, Humboldt-Universitat zu Berlin, Berlin-Mitte, Germany.
SO PROTEINS, **(1996 Jun)** 25 (2) 157-68.
Journal code: 8700181. ISSN: 0887-3585.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199612
ED Entered STN: 19970128
Last Updated on STN: 19970128
Entered Medline: 19961205
AB The predictive limits of the amino acid composition for the secondary structural content (percentage of residues in the secondary structural states helix, sheet, and coil) in proteins are assessed quantitatively. For the first time, techniques for prediction of secondary structural content are presented which rely on the amino acid composition as the only information on the query protein. In our first method, the amino acid composition of an unknown protein is represented by the best (in a least square sense) linear combination of the characteristic amino acid compositions of the three secondary structural types computed from a learning set of tertiary structures. The second technique is a generalization of the first one and takes into account also possible compositional couplings between any two sorts of amino acids. Its mathematical formulation results in an eigenvalue/eigenvector problem of the second moment matrix describing the amino acid compositional fluctuations of secondary structural types in various proteins of a learning set. Possible correlations of the principal directions of the eigenspaces with physical properties of the amino acids were also checked. For example, the first two eigenvectors of the helical eigenspace correlate with the size and hydrophobicity of the residue types respectively. As learning and test sets of tertiary structures, we utilized representative, automatically generated subsets of Protein Data Bank (PDB) consisting of non-homologous protein structures at the resolution thresholds < or = 1.8A, < or = 2.0A, < or = 2.5A, and < or = 3.0 A. We show that the consideration of compositional couplings improves prediction accuracy, albeit not dramatically. Whereas in the self-consistency test (learning with the protein to be predicted), a clear

decrease of prediction accuracy with worsening resolution is observed, the jackknife test (leave the predicted protein out) yielded best results for the largest dataset (< or = 3.0A, almost no difference to the self-consistency test!), i.e., only this set, with more than 400 proteins, is sufficient for stable computation of the parameters in the prediction function of the second method. The average absolute error in predicting the fraction of helix, sheet, and coil from amino acid composition of the query protein are 13.7, 12.6, and 11.4%, respectively with r.m.s. deviations in the range of 8.6 divided by 11.8% for the 3.0 A dataset in a jackknife test. The absolute precision of the average absolute errors is in the range of 1 divided by 3% as measured for other representative subsets of the PDB. Secondary structural content prediction methods found in the literature have been clustered in accordance with their prediction accuracies. To our surprise, much more complex secondary structure prediction methods utilized for the same purpose of secondary structural content prediction achieve prediction accuracies very similar to those of the present analytic techniques, implying that all the information beyond the amino acid composition is, in fact, mainly utilized for positioning the secondary structural state in the sequence but not for determination of the overall number of residues in a secondary structural type. This result implies that higher prediction accuracies cannot be achieved relying solely on the amino acid composition of an unknown query protein as prediction input. Our prediction program SSCP has been made available as a World Wide Web and E-mail service.

CT Check Tags: Comparative Study; Support, Non-U.S. Gov't
*Amino Acids: CH, chemistry
Computer Simulation
Forecasting: MT, methods
Models, Chemical
***Protein Structure, Secondary**
Reproducibility of Results

CN 0 (Amino Acids)

L125 ANSWER 16 OF 19 MEDLINE
AN **96190961** MEDLINE
DN **96190961** PubMed ID: **8609611**
TI The automatic search for ligand binding sites in proteins of known three-dimensional structure using only geometric criteria.
AU Peters K P; Fauck J; **Frommel C**
CS Humboldt-University of Berlin, Medical Faculty (Charite), Institute of Biochemistry, Germany.
SO JOURNAL OF MOLECULAR BIOLOGY, (1996 Feb 16) 256 (1) 201-13.
Journal code: 2985088R. ISSN: 0022-2836.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199605
ED Entered STN: 19960605
Last Updated on STN: 19960605
Entered Medline: 19960524

AB The biological function of a protein typically depends on the structure of specific binding sites. These sites are located at the surface of the protein molecule and are determined by geometrical arrangements and physico-chemical properties of tens of non-hydrogen atoms. In this paper we describe a new algorithm called APROPOS, based purely on geometric criteria for identifying such binding sites using atomic co-ordinates. For the description of the protein shape we use an alpha-shape algorithm which generates a whole family of shapes with different levels of detail. Comparing shapes of different resolution we find cavities on the surface of the protein responsible for ligand binding. The algorithm correctly locates more than 95% of all binding sites for ligands and prosthetic groups of molecular mass between about 100 and 2000 Da in a representative

set of proteins. Only in very few proteins does the method find binding sites of single ions outside the active site of enzymes. With one exception, we observe that interfaces between subunits show different geometric features compared to binding sites of ligands. Our results clearly support the view that protein-protein interactions occur between flat areas of protein surface whereas specific interactions of smaller ligands take place in pockets in the surface.

CT Check Tags: Animal; In Vitro; Support, Non-U.S. Gov't
Algorithms
Amino Acid Sequence
Binding Sites
Carrier Proteins: CH, chemistry
Carrier Proteins: GE, genetics
Carrier Proteins: ME, metabolism
DNA-Binding Proteins: CH, chemistry
DNA-Binding Proteins: GE, genetics
DNA-Binding Proteins: ME, metabolism
Databases, Factual
Heat-Shock Proteins: CH, chemistry
Heat-Shock Proteins: GE, genetics
Heat-Shock Proteins: ME, metabolism
Ligands
Models, Molecular
Molecular Sequence Data
Molecular Structure
Molecular Weight
Protein Binding
Protein Conformation
*Proteins: CH, chemistry
Proteins: GE, genetics
Proteins: ME, metabolism
Pyrophosphatases: CH, chemistry
Pyrophosphatases: GE, genetics
Pyrophosphatases: ME, metabolism
Subtilisins: CH, chemistry
Subtilisins: GE, genetics
Subtilisins: ME, metabolism
Tacrolimus: ME, metabolism
Tacrolimus Binding Proteins

RN 109581-93-3 (Tacrolimus)

CN 0 (Carrier Proteins); 0 (DNA-Binding Proteins); 0 (Heat-Shock Proteins); 0 (Ligands); 0 (Proteins); EC 3.4.21.- (Subtilisins); EC 3.6.1.- (Pyrophosphatases); EC 3.6.1.1 (inorganic pyrophosphatase); EC 5.2.1.- (Tacrolimus Binding Proteins)

L125 ANSWER 17 OF 19 MEDLINE

AN **96180020** MEDLINE

DN **96180020** PubMed ID: **8609628**

TI An evolutionary trace method defines binding surfaces common to protein families.

AU Lichtarge O; Bourne H R; Cohen F E

CS Department of Cellular and Molecular Pharmacology, University of California San Franciso, 94143-0450, USA.

SO JOURNAL OF MOLECULAR BIOLOGY, **(1996 Mar 29)** 257 (2) 342-58.
Journal code: 2985088R. ISSN: 0022-2836.

CY ENGLAND: United Kingdom

DT Journal; Article; (JOURNAL ARTICLE)

LA English

FS Priority Journals; Space Life Sciences

EM 199605

ED Entered STN: 19960605
Last Updated on STN: 19960605
Entered Medline: 19960530

AB X-ray or NMR structures of proteins are often derived without their **ligands**, and even when the structure of a full complex is available, the area of contact that is functionally and energetically significant may be a specialized subset of the geometric interface deduced from the spatial proximity between **ligands**. Thus, even after a structure is solved, it remains a major theoretical and experimental goal to localize protein functional interfaces and understand the role of their constituent residues. The evolutionary trace method is a systematic, transparent and novel predictive technique that identifies active sites and functional interfaces in proteins with known structure. It is based on the extraction of functionally important residues from sequence conservation patterns in homologous proteins, and on their mapping onto the protein surface to generate clusters identifying functional interfaces. The SH2 and SH3 modular signaling domains and the DNA binding domain of the nuclear hormone receptors provide tests for the accuracy and validity of our method. In each case, the evolutionary trace delineates the functional epitope and identifies residues critical to binding specificity. Based on mutational evolutionary analysis and on the structural homology of protein families, this simple and versatile approach should help focus site-directed mutagenesis studies of structure-function relationships in macromolecules, as well as studies of specificity in molecular recognition. More generally, it provides an evolutionary perspective for judging the functional or structural role of each residue in protein structure.

CT Check Tags: Animal; Human; Support, Non-U.S. Gov't; Support, U.S. Gov't, P.H.S.
Amino Acid Sequence
***Binding Sites**
Conserved Sequence
DNA-Binding Proteins: CH, chemistry
Databases, Factual
*Evolution, Molecular
Models, Molecular
Molecular Sequence Data
Mutation
Protein Conformation
***Proteins: CH, chemistry**
Rats
Receptors, Glucocorticoid: CH, chemistry
Sequence Alignment: MT, methods
Sequence Homology, Amino Acid
Zinc Fingers
src Homology Domains

CN 0 (DNA-Binding Proteins); 0 (Proteins); 0 (Receptors, Glucocorticoid)

L125 ANSWER 18 OF 19 MEDLINE
AN **96038917** MEDLINE
DN **96038917** PubMed ID: **7584434**
TI Constituting a receptor-**ligand** information base from quality-enriched data.
AU Hemm K; Aberer K; Hendlich M
CS GMD-IPSI, Darmstadt, Germany.
SO ISMB, **(1995)** 3 170-8.
Journal code: 9509125.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199512
ED Entered STN: 19960124
Last Updated on STN: 19960124
Entered Medline: 19951205
AB Many different resources are needed for analyzing relevant experimental

data in drug design. Currently this data is difficult to access, because it is stored in heterogeneous databases, spread over many platforms, poorly interconnected, incomplete, erroneous, or just not electronically available. In order to establish a high quality database for drug design we have developed a new demand-driven methodology for integrating and semantically enriching heterogeneous data from different research areas and for migrating the data into an object-oriented database management system. In this way we have established a database containing well-prepared, relevant data needed for drug design and offering the advantages of modern database technology, like a comprehensive object-oriented data model, a flexible declarative query language and support for persistent storage and sharing of data in a multi-user environment.

CT Check Tags: Comparative Study; Support, Non-U.S. Gov't
Amino Acid Sequence
Binding Sites
***Computer Simulation**
***Databases, Factual**
Drug Design
***Ligands**
Molecular Sequence Data
Protein Structure, Secondary
***Proteins: CH, chemistry**
***Proteins: ME, metabolism**
***Receptors, Cell Surface: CH, chemistry**
***Receptors, Cell Surface: ME, metabolism**
***Receptors, Drug: CH, chemistry**
***Receptors, Drug: ME, metabolism**
Sequence Homology, Amino Acid

CN 0 (**Ligands**); 0 (Proteins); 0 (Receptors, Cell Surface); 0 (Receptors, Drug)

L125 ANSWER 19 OF 19 MEDLINE
AN **95327685** MEDLINE
DN **95327685** PubMed ID: **7604031**
TI Alternate protein frameworks for molecular recognition.
AU Ku J; Schultz P G
CS Howard Hughes Medical Institute, University of California, Berkeley 94720, USA.
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, **(1995 Jul 3)** 92 (14) 6552-6.
Journal code: 7505876. ISSN: 0027-8424.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 199508
ED Entered STN: 19950822
Last Updated on STN: 19980206
Entered Medline: 19950810
AB In an effort to determine whether proteins with structures other than the immunoglobulin fold can be used to mimic the **ligand** binding properties of antibodies, we generated a library from the four-helix bundle protein cytochrome b562 in which the two loops were randomized. Panning of this library against the bovine serum albumin (BSA) conjugate of N-methyl-p-nitrobenzylamine derivative 1 by phage display methods yielded cytochromes in which residues Trp-20, Arg-21, and Ser-22 in loop A and Arg-83 and Trp-84 in loop B were conserved. The individual mutants, which fold into native-like structure, bind selectively to the BSA-1 conjugate with micromolar dissociation constants (Kd), in comparison to a monoclonal antibody that binds selectively to 1 with a Kd of 290 nM. These and other antibody-like receptors may prove useful as therapeutic agents or as reagents for both intra- and extracellular studies.